L'Electron Magnétique

L'ÉLECTRON MAGNÉTIQUE

(THÉORIE DE DIRAC)

PAR

LOUIS DE BROGLIE

MEMBRE DE L'INSTITUT

PROFESSEUR A LA SORBONNE

LAUREAT DU PRIX NOBEL

PARIS

HERMANN ET C^{IE}, ÉDITEURS

6, RUE DE LA SORBONNE, 6

—

1934

PRÉFACE

La théorie de l'électron due à M. Dirac présente un grand intérêt à plusieurs points de vue. Elle constitue la forme la plus parfaite que nous possédions actuellement de la dynamique ondulatoire de l'électron ; elle concilie, au moins dans une certaine mesure, les idées relativistes et les conceptions quantiques ; elle précise sous une forme en accord avec les principes de la nouvelle Physique l'hypothèse si féconde du corpuscule électrisé, magnétique et tournant de MM. Uhlenbeck et Goudsmit ; elle permet enfin, tant en ce qui concerne la structure fine des spectres qu'en ce qui concerne les effets Zeeman anormaux, de rendre compte d'importants faits expérimentaux et, par là, elle reçoit à son tour une belle confirmation. Il nous a paru utile d'en publier un exposé général en rédigeant des cours que nous avons faits dans les plus récentes années à l'Institut Henri Poincaré.

Pour bien montrer que la théorie de Dirac n'est pas un simple jeu de théoricien, mais qu'elle vient à point pour nous permettre d'expliquer des faits importants, nous avons cru devoir faire dans la première partie du livre un exposé des phénomènes qui ont reçu d'elle une interprétation satisfaisante en montrant en même temps que ces phénomènes étaient rebelles à toute explication complète par les anciennes

théories quantiques ou même par la Mécanique ondulatoire sous sa forme initiale.

Nous avons réservé aussi une certaine place dans cette première partie de l'ouvrage au rappel des principes généraux de la nouvelle Mécanique et de la conception des lois physiques qui en découle; il est impossible, en effet, de bien comprendre la théorie de Dirac si l'on n'a pas ces principes présents à l'esprit.

Dans l'exposé de la théorie qui fait l'objet de la seconde partie, nous avons conservé la forme un peu dissymétrique des équations employée dès le début par M. Dirac lui-même sans chercher à prendre des notations plus symétriques au point de vue relativiste comme l'ont fait nombre d'auteurs par la suite. Cette recherche d'une symétrie de forme nous paraît, en effet, un peu vaine puisque, comme nous avons cherché à le montrer dans le dernier chapitre, la théorie de Dirac, malgré l'invariance de forme de ses équations pour une transformation de Lorentz, doit faire jouer au temps un rôle particulier pour rester en accord avec les principes généraux actuels de la Physique quantique. A la fin de la seconde partie, nous avons consacré un chapitre à une vue systématique de l'ensemble de la théorie qui pourra aider le lecteur à en saisir l'harmonie.

La troisième partie de l'ouvrage comprend l'interprétation par la théorie de Dirac des faits expérimentaux rappelés dans la première partie. Puis vient une étude de certaines conséquences un peu étranges des équations fondamentales, en particulier la prévision des états à énergie négative. Nous avons exposé ces difficultés sans en préconiser aucune solution; quelle que soit la manière dont on sera amené dans l'avenir à les résoudre, elles méritent d'être étudiées car elles ont leurs racines dans les bases même de la théorie.

Nous espérons que l'ensemble de l'ouvrage permettra au lecteur de saisir à la fois la beauté de la théorie de Dirac, son utilité pour l'interprétation des faits expérimentaux et aussi ses lacunes et ses points faibles.

J'exprime toute ma reconnaissance à M. Jean-Louis Destouches pour m'avoir aidé dans la correction des épreuves.

Louis de Broglie.

PREMIÈRE PARTIE

SUCCÈS ET ÉCHECS DES THÉORIES QUANTIQUES ET DE LA MÉCANIQUE ONDULATOIRE PRIMITIVE

CHAPITRE PREMIER

Le spectre atomique de l'hydrogène. Théories de Bohr et Sommerfeld

1. *La formule de Balmer et les termes spectraux de l'hydrogène*

La série de raies la plus anciennement connue dans le spectre visible de l'hydrogène est la série de Balmer. Elle est formée de 4 raies principales (qui sont en réalité de petits doublets comme nous le verrons plus tard). Voici les noms et les longueurs d'onde de ces 4 raies :

H_α : 6.563 Å H_β : 4.861 Å H_γ : 4.340 Å H_δ : 4.102 Å

Balmer est parvenu, il y a déjà un demi-siècle, à trouver une formule qui donne les fréquences de ces raies. Cette formule est la suivante :

$$\nu_m = R\left[\frac{1}{4} - \frac{1}{m^2}\right] \qquad m = 3, 4, 5, 6 \qquad (1)$$

ν_3 est la fréquence de la raie H_α, ν_4 celle de H_β, etc. R est une constante dite « constante de Rydberg » très sensiblement égale à $3,29201 . 10^{15}$.

D'autres séries de raies découvertes postérieurement dans le spectre non visible de l'hydrogène suivent des lois ana-

logues. Telles sont notamment la série ultra-violette de Lyman pour laquelle la fréquence des raies est donnée par la formule :

$$\nu_m = R\left[1 - \frac{1}{m^2}\right] \qquad m = 2, 3... \qquad (2)$$

et la série infra-rouge de Paschen pour laquelle on a :

$$\nu_m = R\left[\frac{1}{9} - \frac{1}{m^2}\right] \qquad m = 4, 5... \qquad (3)$$

On voit que toutes ces formules rentrent dans le type général :

$$\nu = R\left[\frac{1}{n^2} - \frac{1}{m^2}\right] \qquad n < m \qquad (4)$$

Pour les raies de la série de Lyman, $n = 1$; pour celles de la série de Balmer, $n = 2$; pour celles de la série de Paschen, $n = 3$.

De cette constatation, on peut s'élever à une loi générale qui s'est montrée exacte pour toutes les raies spectrales de tous les corps. C'est le « principe de combinaison de Ritz » dont voici l'énoncé :

« La fréquence de toute raie spectrale est égale à la différence de deux termes spectraux caractéristiques du corps émetteur. »

Ou sous une autre forme :

« Pour tout corps émetteur, on peut dresser un tableau de nombres dits termes spectraux tels que la fréquence de toute raie spectrale du corps soit la différence de deux de ses termes spectraux. »

La formule (4) nous montre alors que, pour l'hydrogène, les termes spectraux ont, au moins en valeur absolue, la forme $\frac{R}{n^2}$ avec $n = 1, 2, \ldots$

Une étude expérimentale plus perfectionnée des raies de la série de Balmer a d'ailleurs montré que ces raies sont en réalité chacune formées de deux raies très voisines. Autrement dit, quand on analyse avec une dispersion suffisante

la série de Balmer, on s'aperçoit que chacune des raies considérées primitivement comme simple est en réalité un petit doublet. Pour chacun de ces doublets la différence de fréquence entre les deux composantes est la même. Nous verrons plus loin comment Sommerfeld a pu interpréter cette structure fine de la série de Balmer.

2. ***La théorie de Bohr pour les termes spectraux de l'hydrogène.*** — En 1912, M. Bohr est parvenu à interpréter les termes spectraux de l'hydrogène et, ce faisant, il a fondé sur des bases nouvelles la théorie moderne de l'atome.

Un peu avant la théorie de Bohr, les physiciens, non sans avoir quelque peu tâtonné, avaient fini, suivant la suggestion de lord Rutherford, par adopter un modèle planétaire de l'atome. Selon cette vue, l'atome du corps simple dont le rang dans la série de Mendeleeff est N serait muni d'un noyau central de charge positive Ne, e étant la charge électrique élémentaire $+4{,}77 \,.\, 10^{-10}$ u.e.s.; autour de ce noyau, tourneraient N électrons de charge $-e$ de sorte que l'ensemble de l'atome serait électriquement neutre. M. Bohr a eu l'idée de soumettre au calcul ce modèle d'atome en lui appliquant les lois de quanta introduites par Planck avec succès dans l'étude du rayonnement noir. Il admet que dans l'atome un électron-planète ne peut décrire autour du soleil positif central que certains des mouvements prévus par la Mécanique classique. A ces mouvements stables des électrons correspondent pour l'atome des « états stationnaires » durant lesquels, contrairement aux prévisions de la théorie électromagnétique classique, aucun rayonnement n'est émis. L'émission des raies spectrales ne peut donc avoir lieu que lors du passage brusque de l'atome d'un état stationnaire initial à un autre état stationnaire de moindre énergie. Quelle sera la fréquence de la raie émise pendant un tel passage? Bohr la détermine en admettant que l'énergie perdue par l'atome est rayonnée sous forme d'un seul quantum de lumière de valeur $h\nu$, sous la forme d'un seul photon pour employer le langage actuel. Si donc E_i et E_j désignent les

énergies de l'atome dans l'état stationnaire initial et dans l'état stationnaire final, la fréquence ν_{ij} émise lors de la transition est :

$$\nu_{ij} = \frac{E_i - E_j}{h} \tag{5}$$

Cette formule explique immédiatement le principe de combinaison de Ritz et nous montre que les termes spectraux d'un atome sont égaux aux énergies de ses divers états stationnaires divisées par la constante de Planck.

Le problème essentiel est donc de déterminer les énergies des états stationnaires. Pour cela, Bohr dans son travail initial admet toujours que l'électron se comporte comme une charge ponctuelle obéissant aux lois de la Dynamique de Newton, mais il restreint le nombre des mouvements possibles en introduisant les règles de quantification de Planck. Au moment où Bohr écrivait, on ne savait quantifier que les mouvements périodiques définis par une seule variable q. La méthode de quantification pour ce cas était la suivante : p étant le moment de Lagrange conjugué de la variable q, on écrivait :

$$\oint p\, d\, q = n\, h \qquad (n \text{ entier}) \tag{6}$$

l'intégrale étant étendue à un cycle entier du mouvement et h étant la constante de Planck. Bohr a été amené naturellement à supposer que les mouvements de l'électron satisfaisant à la condition (6) sont les mouvements stables correspondant aux états stationnaires de l'atome.

La méthode permet de calculer aisément l'énergie des trajectoires circulaires stables de l'atome d'hydrogène qui, dans la conception de Rutherford, est formé d'un noyau de charge $+e$ et d'un électron-planète de charge $-e$. Si θ est l'azimut qui repère la position de l'électron sur une trajectoire circulaire de rayon r, la condition (6) donne :

$$m\, r^2\, \dot{\theta} = n \cdot \frac{h}{2\pi} \qquad \left(\dot{\theta} = \frac{d\,\theta}{d\,t}\right) \tag{7}$$

Elle exprime que le moment de rotation de l'électron sur l'orbite stable est un multiple entier de $\frac{h}{2\pi}$ Comme, d'autre part, les lois de la Dynamique fournissent la relation :

$$m r \dot{\theta}^2 = \frac{e^2}{r^2} \tag{8}$$

on trouve aisément pour l'énergie du n^e mouvement circulaire quantifié :

$$E_n = \frac{1}{2} m r^2 \dot{\theta}^2 - \frac{e^2}{r} = -\frac{1}{2} m r^2 \dot{\theta}^2 = -\frac{2\pi^2 m e^4}{n^2 h^2} \tag{9}$$

Donc, si l'on s'en tient aux mouvements circulaires, les termes spectraux de l'hydrogène doivent être de la forme :

$$\frac{E_n}{h} = -\frac{2\pi^2 m e^4}{n^2 h^3} \tag{10}$$

D'après la formule (5), les raies de l'hydrogène doivent avoir des fréquences données par la relation générale :

$$\nu_{n n'} = \frac{2\pi^2 m e^4}{h^3}\left(\frac{1}{n^2} - \frac{1}{n'^2}\right) \qquad (n' > n) \tag{11}$$

et l'on retombe sur la formule (4) tirée de l'expérience en posant :

$$R = \frac{2\pi^2 m e^4}{h^3} \tag{12}$$

Or le calcul numérique du second membre de (12) montre que sa valeur est bien sensiblement égale à la valeur expérimentale de la constante de Rydberg.

On peut recommencer le même calcul en supposant que l'on a affaire à un atome de rang atomique N ionisé N — 1 fois. On a alors à traiter un problème identique à celui de l'atome d'hydrogène sauf que la charge centrale est Ne au lieu de e [1]. En reprenant le calcul, on trouve aisément pour les termes spectraux à la place de (10) :

[1] On dit souvent que l'atome de rang N ionisé (N — 1) fois est hydrogénoïde.

$$\frac{E_n}{h} = -\frac{2\pi^2 m e^4}{n^2 h^3} N^2 = -\frac{RN^2}{n^2} \qquad (13)$$

Les termes spectraux sont multipliés par le carré du nombre atomique. Le cas le plus simple est celui de l'hélium une fois ionisé pour lequel on a $N = 2$. Les termes spectraux et les fréquences sont quadruplés. Les nombres expérimentaux indiquent cependant que la constante de Rydberg n'a pas tout à fait la même valeur pour H et pour He^+. M. Bohr a pu rendre compte de cette différence en faisant intervenir la réaction de l'électron sur le noyau.

3. *Energies quantifiées des orbites elliptiques.* — Les calculs de Bohr rappelés ci-dessus ne peuvent être regardés comme complets car l'étude du mouvement d'un électron autour d'un noyau est en principe un problème à deux variables : rayon-vecteur et azimut. Eliminer les variations du rayon-vecteur en ne considérant que les trajectoires circulaires est évidemment un procédé artificiel. Mais pour traiter complètement le problème, il faut d'abord savoir écrire les conditions de quanta pour les mouvements à plusieurs degrés de liberté. Voici comment on y est parvenu.

Soit un système à n degrés de liberté défini par n variables $q_1 \ldots q_n$. Si toutes les variables admettent une même période de variation T, c'est-à-dire si à des intervalles de temps égaux à T elles reprennent la même valeur, le système reprend régulièrement les mêmes configurations : il est périodique. Si chaque variable q_i a une période T_i de variation, ces périodes étant incommensurables entre elles, le système est quasi-périodique. Les conditions de quanta n'ont de sens que pour les systèmes périodiques ou quasi-périodiques. Pour les systèmes de ce genre que l'on a eu à quantifier dans l'ancienne théorie des quanta, il a toujours été possible de choisir les variables de façon qu'elles forment un système de « variables séparées » c'est-à-dire que chacun des moments de Lagrange p_i puisse s'exprimer en fonction de la seule coordonnée q_i correspondante. Les variables étant ainsi choisies, MM. H. A.

Wilson et Sommerfeld ont montré que la quantification doit s'exprimer par les n conditions :

$$\oint p_i \, dq_i = n_i \, h \qquad (n_i \text{ entier}) \quad i = 1, 2, \dots n \,. \tag{14}$$

chaque intégrale étant prise pour la période T_i de la variable q_i.

M. Sommerfeld s'est servi de ce nouvel énoncé pour traiter plus complètement le problème de l'atome d'hydrogène en tenant compte de tous les mouvements elliptiques. Soit r le rayon vecteur et θ l'azimut de l'électron sur sa trajectoire képlérienne. L'énergie cinétique T sera :

$$T = \frac{1}{2} m \,(\dot{r}^2 + r^2 \dot{\theta}^2) \tag{15}$$

et les moments de Lagrange seront par définition :

$$p_r = \frac{\partial T}{\partial \dot{r}} = m \dot{r} \qquad p_\theta = \frac{\partial T}{\partial \dot{\theta}} = m r^2 \dot{\theta} \tag{16}$$

Les conditions (14) peuvent ici s'appliquer et s'écrivent :

$$\int_0^{2\pi} m r^2 \dot{\theta} \, d\theta = n_1 h \qquad \oint m \dot{r} \, dr = n_2 h \tag{17}$$

Or $m r^2 \dot{\theta}$ est le moment de rotation constant dans un champ central d'après le théorème des aires. La première condition (17) donne donc :

$$m r^2 \dot{\theta} = n_1 \frac{h}{2\pi} \tag{18}$$

et coïncide avec la condition (7) de Bohr pour les orbites circulaires.

Pour calculer la deuxième intégrale (17), nous devons écrire l'expression de l'énergie :

$$E = \frac{1}{2} m \,(\dot{r}^2 + r^2\dot{\theta}^2) - \frac{e^2}{r} = \frac{1}{2m} \left[\frac{(m r^2 \dot{\theta})^2}{r^2} + m^2 \dot{r}^2 \right] - \frac{e^2}{r} \tag{19}$$

d'où l'on tire en tenant compte de (18) :

$$m \dot{r} = p_r = \pm \sqrt{2m \left(E + \frac{e^2}{r} \right) - \frac{n_1^2 h^2}{4 \pi^2 r^2}} \tag{20}$$

formule qui montre la séparation des variables.

Au cours du mouvement, le rayon vecteur r oscille entre les valeurs r_1 et r_2 qui annulent le radical (20) car p_r doit être réel. En supposant $r_1 < r_2$, on écrira :

$$\oint p_r dr = 2 \int_{r_1}^{r_2} \sqrt{2m\left(\mathrm{E} + \frac{e^2}{r}\right) - \frac{n_1^2 h^2}{4\pi^2 r^2}}\, dr \qquad (21)$$

car on doit prendre dans (20) le signe + quand r augmente et le signe — quand il diminue. Sommerfeld a calculé l'intégrale (21) par la méthode des résidus de Cauchy et lui a trouvé la valeur $-|n_1| h + \frac{2\pi m e^2}{\sqrt{2m|\mathrm{E}|}}$. En égalant à $n_2 h$, on trouve aisément :

$$\mathrm{E}_{n_1, n_2} = -\frac{2\pi^2 m e^4}{(|n_1| + n_2)^2 h^2} \qquad (22)$$

Ceci donne l'énergie quantifiée de l'état stationnaire correspondant aux nombres quantiques n_1 et n_2. Comme $|n_1|$ et n_2 sont des nombres entiers positifs ou nuls qui ne peuvent être tous deux nuls ainsi qu'on le voit aisément, on peut poser :

$$|n_1| + n_2 = n \qquad (n = 1, 2 \ldots\ldots) \qquad (33)$$

et la formule (22) donne les mêmes niveaux d'énergie que la théorie primitive de Bohr. Autrement dit, le fait de considérer les orbites elliptiques ne conduit à aucun terme spectral nouveau. L'introduction de deux degrés de liberté ne peut expliquer à elle seule la structure fine de la série de Balmer.

4. *Théorie de la structure fine de Sommerfeld.* — Pour expliquer la structure fine du spectre de l'hydrogène, Sommerfeld a eu l'idée de chercher à employer la Mécanique de Relativité au lieu de la Mécanique classique. Ceci se justifie en remarquant que, dans l'atome de Bohr, la vitesse des électrons sur les orbites internes doit être comparable à la vitesse de lumière.

Dans la dynamique relativiste, l'énergie cinétique de l'électron a pour expression :

$$\mathrm{T} = m_0 c^2 \left[\frac{1}{\sqrt{1-\beta^2}} - 1\right] \qquad (24)$$

m_0 étant la masse de l'électron au repos et β ayant la signification usuelle :

$$\beta = \frac{v}{c} = \frac{1}{c}\sqrt{\dot{r}^2 + r^2\dot{\theta}^2} \tag{25}$$

Mais ici on ne doit plus définir le moment p_i conjugué de la variable q_i comme étant égal à la dérivée de T par rapport à $\dot{q}_i$; on doit introduire la fonction de Lagrange relativiste :

$$L = -m_0 c^2\sqrt{1-\beta^2} - U \tag{26}$$

U étant l'énergie potentielle et poser la définition :

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \tag{27}$$

Dans le problème de l'atome d'hydrogène, l'énergie potentielle ne dépend pas des q_i et l'on a :

$$\begin{aligned} p_r &= \frac{\partial L}{\partial \dot{r}} = \frac{m_0 c^2}{2\sqrt{1-\beta^2}} \cdot \frac{\partial \beta^2}{\partial \dot{r}} = \frac{m_0 \dot{r}}{\sqrt{1-\beta^2}} \\ p_\theta &= \frac{\partial L}{\partial \dot{\theta}} = \frac{m_0 c^2}{2\sqrt{1-\beta^2}} \cdot \frac{\partial \beta^2}{\partial \dot{\theta}} = \frac{m_0 r^2 \dot{\theta}}{\sqrt{1-\beta^2}} \end{aligned} \tag{28}$$

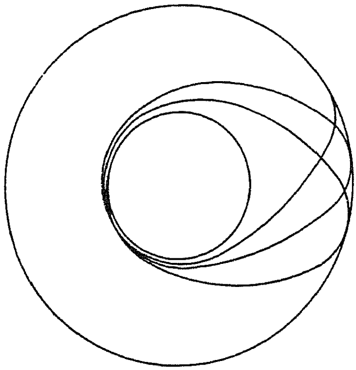
Les conditions de quanta (14) sont donc ici :

$$\int_0^{2\pi} \frac{m_0 r^2 \dot{\theta}}{\sqrt{1-\beta^2}}\, d\theta = n_1 h \; ; \qquad \oint \frac{m_0 \dot{r}}{\sqrt{1-\beta^2}}\, dr = n_2 h \tag{29}$$

L'étude détaillée de la trajectoire que nous ne ferons pas ici montre que l'électron décrit une ellipse avec rotation du périhélie ; autrement dit la trajectoire est à chaque instant tangente à une ellipse osculatrice qui tourne lentement dans son plan. Le rayon vecteur oscille entre deux valeurs r_1 et r_2, mais le temps qu'il met à décrire le cycle $r_1 \longrightarrow r_2 \longrightarrow r_1$ (période de la variable r) est un peu plus long que le temps mis par l'azimut pour augmenter de 2π (période de la variable θ). L'orbite ne se ferme donc pas exactement et le mouvement est quasi-périodique.

Le moment p_θ est encore le moment de rotation autour

centre et on démontre aisément qu'il est encore ici constant, c'est-à-dire que le théorème des aires est toujours valable. La première condition (29) donne donc :


FIG. 1

$$p_\theta = \frac{m_0 r^2 \dot{\theta}}{\sqrt{1-\beta^2}} = n_1 \frac{h}{2\pi} \qquad (30)$$

L'énergie totale est la somme de l'énergie interne $m_0 c^2$ de l'énergie cinétique et de l'énergie potentielle; elle est donc égale à :

$$W = \frac{m_0 c^2}{\sqrt{1-\beta^2}} - \frac{e^2}{r} \qquad (31)$$

En tenant compte de (28) et de (25), on vérifie facilement la formule :

$$W = c\sqrt{m_0^2 c^2 + p_r^2 + \frac{p_\theta^2}{r^2}} - \frac{e^2}{r} \qquad (32)$$

Appelons E l'énergie W diminuée du terme $m_0 c^2$; E est alors l'énergie telle qu'elle se définit en Mécanique classique. Remplaçons dans (32) W par $E + m_0 c^2$ et résolvons par rapport à p_r. Il vient :

$$p_r = \pm\sqrt{A + 2\frac{B}{r} + \frac{C}{r^2}} \qquad (33)$$

avec les notations :

$$A = \frac{E^2}{c^2} + 2 m_0 E = m_0^2 c^2 \left[\left(1 + \frac{E}{m_0 c^2}\right)^2 - 1\right]$$

$$B = \frac{E e^2}{c^2} + m_0 e^2 E \qquad (34)$$

$$C = \frac{e^4}{c^2} - p_\theta^2 = \frac{e^4}{c^2} - \frac{n_1^2 h^2}{4\pi^2}$$

Sommerfeld introduit alors dans le calcul ce qu'on appelle « la constante de structure fine » :

$$\alpha = \frac{2\pi e^2}{hc} \qquad (\alpha^2 = 5{,}2 . 10^{-5}\ \text{c.g.s.}) \tag{35}$$

On peut donc écrire :

$$C = -\frac{n_1{}^2 h^2}{4\pi^2}\left[1 - \frac{\alpha^2}{n_1{}^2}\right] \tag{36}$$

Une application du théorème des résidus permet d'établir que l'on a :

$$\oint\sqrt{A + \frac{2B}{r} + \frac{C}{r^2}}\,dr = -2\pi i\left(\sqrt{C} - \frac{B}{\sqrt{A}}\right) \tag{37}$$

En égalant le second membre de (37) à $n_2 h$ conformément à (29) et en remplaçant A, B et C par leurs valeurs, Sommerfeld trouve après un calcul simple :

$$1 + \frac{E}{m_0 c^2} = \left[1 + \frac{\alpha^2}{\left[n_2 + \sqrt{n_1{}^2 - \alpha^2}\right]^2}\right]^{-\frac{1}{2}} \tag{38}$$

formule qui donne rigoureusement l'énergie E de l'état stationnaire défini par les nombres quantiques n_1 et n_2.

La quantité α^2 étant très petite, une première approximation consiste à négliger les termes de degré supérieur au premier en α^2. On retombe alors, comme on devait s'y attendre, sur la formule (22) de la théorie non relativiste et on ne trouve pas la structure fine. Une approximation meilleure consistera à garder les termes en α^4 et à écrire :

$$\left[1 + \frac{\alpha^2}{\left[n_2{}^2 + \sqrt{n_1{}^2 - \alpha^2}\right]^2}\right]^{-\frac{1}{2}} =$$

$$1 - \frac{\alpha^2}{2(|n_1| + n_2)^2} - \frac{1}{2}\,\frac{\alpha^4}{(|n_1| + n_2)^4}\left(\frac{1}{4} + \frac{n_2}{|n_1|}\right) \tag{39}$$

En portant cette valeur dans (38), on trouve :

$$E = -\frac{2\pi^2 m_0 e^4}{(|n_1| + n_2)^2 h^2}\left[1 + \frac{\alpha^2}{(|n_1| + n_2)^2}\left(\frac{1}{4} + \frac{n_2}{|n_1|}\right)\right] \tag{40}$$

Le troisième terme entre crochets explique l'existence d'une structure fine de la série de Balmer car il dépend séparément de $|n_1|$ et n_2 et non pas seulement de $|n_1| + n_2$.

Nous allons maintenant changer un peu les notations. Nous appellerons « nombre quantique total » le nombre $n=|n_1|+n_2$ et « nombre quantique azimutal » le nombre $k=|n_1|$. Il est évident que l'on peut caractériser chaque niveau quantifié d'énergie par les nombres n et k au lieu des nombres n_1 et n_2. La formule (40) s'écrit maintenant (en introduisant la constante de Rydberg) :

$$E_{n,k}=-\frac{Rh}{n^2}\left[1+\frac{\alpha^2}{n^2}\left(\frac{n}{k}-\frac{3}{4}\right)\right] \tag{41}$$

Dans l'ancienne théorie des quanta, on admettait que le nombre azimutal k ne pouvait jamais prendre la valeur 0 ; pour les trajectoires circulaires on avait $n_2=0$ ou $n=k$ et pour les trajectoires elliptiques $0<k<n$. D'après la formule (41), chaque orbite stationnaire est caractérisée par une énergie E_{nk} qui ne dépend plus seulement de n, mais aussi de k ; mais comme α^2 est très petit devant l'unité, les divers termes spectraux correspondant à une même valeur de n sont très voisins et l'on obtient bien ainsi une structure fine des raies prévues par la théorie de Bohr.

Il est visible que le terme spectral de Bohr correspondant à une valeur donnée de n se décompose ici en n termes voisins car pour n fixé on peut avoir pour k les n valeurs 1, 2, ... n. Il convient de remarquer que l'écartement des termes voisins est d'autant plus petit que n est plus élevé en raison de la présence de n_2 au dénominateur du terme en α^2.

Considérons la série de Balmer. En première approximation, les fréquences des raies sont données par la formule :

$$\nu=R\left[\frac{1}{2^2}-\frac{1}{n^2}\right] \qquad n=3, 4 \dots \tag{42}$$

En deuxième approximation, on doit d'après Sommerfeld remplacer le terme spectral $\frac{R}{2^2}$ par $\frac{R}{2^2}\left[1+\frac{\alpha^2}{2^2}\left(\frac{2}{k}-\frac{3}{4}\right)\right]$ et de même le terme spectral $\frac{R}{n^2}$ par $\frac{R}{n^2}\left[1+\frac{\alpha^2}{n^2}\left(\frac{n}{k}-\frac{3}{4}\right)\right]$.

Il y a donc dans la série de Balmer une structure fine à écar-

tement constant provenant du dédoublement ($k=1,2$) du premier terme spectral fixe et une autre structure fine à écartement décroissant quand on s'élève dans la série due à la complexité du deuxième terme spectral variable. Cette deuxième structure fine est pratiquement inobservable parce qu'elle est trop fine. La première correspond à une décomposition de chacune des raies prévues par la formule de Balmer en un doublet à écartement constant pour toute la série et égal à :

$$\Delta \nu_{II} = \frac{R}{2^2}\left[1 + \frac{\alpha}{2^2}\left(\frac{2}{1} - \frac{3}{4}\right)\right] - \frac{R}{2^2}\left[1 + \frac{\alpha^2}{2^2}\left(\frac{2}{2} - \frac{3}{4}\right)\right] = \frac{R\alpha^2}{16} \qquad (43)$$

Le calcul numérique de $\frac{R\alpha^2}{16}$ donne 0,365 cm^{-1} en remplaçant les fréquences par les nombres d'onde. Cette valeur est en assez bon accord avec le nombre expérimental. On obtient donc une interprétation de l'existence de doublets dans la série de Balmer en complétant la théorie de Bohr par l'introduction de la Relativité.

Si l'on reprend les calculs de Sommerfeld en supposant que l'on a affaire non plus à l'atome d'hydrogène mais à un atome de nombre atomique N ionisé ($N-1$) fois, on trouve pour l'énergie de l'état stationnaire caractérisé par les nombres n et k :

$$E_{nk} = -\frac{RN^2 h}{n^2}\left[1 + \frac{\alpha^2 N^2}{n^2}\left(\frac{n}{k} - \frac{3}{4}\right)\right] \qquad (44)$$

La correction de Relativité dans le crochet est donc multipliée par N^2 et l'écartement des doublets par N^4. Pour l'hélium ionisé ($N=2$), l'écartement des doublets dans la série correspondant à la série de Balmer devra donc être 16 fois plus grand que dans la série de Balmer elle-même. On conçoit donc que l'étude des doublets dans le spectre de He^+ ait pu servir à vérifier la théorie de Sommerfeld : cette vérification faite par Paschen a été très satisfaisante.

La théorie de la structure fine de Sommerfeld a donc donné pour H et He^+ de très bons résultats ; elle a pu être étendue avec succès à l'interprétation d'une classe importante de dou-

blets dans le spectre des Rayons X. Mais on s'est ensuite aperçu que pour les raies X et même pour le spectre simple de l'hydrogène, l'application de la formule (41) se heurtait à une grave difficulté. C'est ce que nous expliquerons en détail dans le chapitre III.

CHAPITRE II

Notions sommaires sur les spectres optiques de doublets et leur interprétation

1. *Formule de Rydberg et séries de raies*

LE succès de la formule de Balmer et des formules du même type pour l'hydrogène a conduit depuis longtemps les spectroscopistes à rechercher s'il était possible de trouver des formules analogues pour les éléments autres que l'hydrogène. Le principe de combinaison étant valable pour toutes les raies optiques, il s'agissait de trouver pour les termes spectraux d'un élément quelconque une expression généralisant l'expression $\frac{R}{n^2}$ valable pour l'hydrogène

Rydberg a montré qu'en première approximation les termes spectraux d'un élément peuvent s'écrire sous la forme

$$\frac{R}{(n+\Delta)^2} \qquad (1)$$

où R est la même constante que pour l'hydrogène. Dans l'expression (1) le nombre n est un nombre entier positif et Δ est un nombre non entier susceptible de prendre pour chaque élément plusieurs valeurs différentes. La formule de Rydberg ne représente pas très exactement les termes spec-

traux : Ritz a proposé une expression plus exacte où figurent également le nombre entier n et la quantité Δ. Sans nous occuper ici de l'expression la plus exacte des termes spectraux en fonction de n et de Δ, nous admettrons que chaque terme spectral peut s'exprimer à l'aide de ces deux quantités et peut par suite être représenté par la notation (n, Δ). Pour une valeur donnée de n, Δ est en général susceptible de plusieurs valeurs différentes ; à un même n correspondent donc plusieurs termes spectraux. Les spectroscopistes ont pris l'habitude de désigner les valeurs possibles de Δ par une suite de lettres : s, p, d, f, g, h... et, en dressant le tableau des termes spectraux, ils se sont aperçus que ce tableau a l'aspect suivant :

$$\begin{array}{cccccl} (1, s) & (2, s) & (3, s) & (4, s) & (5, s) & \ldots\ldots\ldots \\ & (2, p) & (3, p) & (4, p) & (5, p) & \ldots\ldots\ldots \\ & & (3, d) & (4, d) & (5, d) & \ldots\ldots\ldots \quad (2) \\ & & & (4, f) & (5, f) & \ldots\ldots\ldots \\ & & & & (5, g) & \ldots\ldots\ldots \end{array}$$

Ainsi pour $n=1$, Δ a la seule valeur s. Pour $n=2$, Δ peut avoir les valeurs s ou p. Pour $n=3$, Δ a une des valeurs s, p, d, etc... D'une façon générale quand n croît d'une unité, le nombre des valeurs possibles de Δ croît aussi d'une unité.

Comme pour l'hydrogène, les fréquences des raies d'une même série spectrale s'obtiennent toujours en faisant la différence du terme constant caractéristique de la série et d'un terme variable. Voici les formules de fréquences de 4 séries qui ont été retrouvées dans tous les spectres et sont particulièrement familières aux spectroscopistes :

Série principale	$\nu_n = (1, s) - (n, p)$	$n = 2,3 \ldots\ldots$
Série diffuse ou 1^re^ série second.	$\nu_n = (2, p) - (n, d)$	$n = 3,4 \ldots\ldots$
Série étroite ou 2^e^ série second.	$\nu_n = (2, p) - (n, s)$	$n = 3,4 \ldots\ldots$ (3)
Série de Bergmann ou série fondamentale	$\nu_n = (3, d) - (n, f)$	$n = 4,5 \ldots\ldots$

Comme les termes spectraux diminuent toujours de valeur quand n augmente, on voit qu'en s'élevant dans une série spectrale, on trouve des raies dont la fréquence se rapproche de plus en plus du terme spectral constant qui caractérise la série ; ce terme spectral peut donc être appelé « la limite de la série ».

Sur les formules de séries (2), on remarque une particularité qui se retrouve pour toutes les séries obtenues dans les conditions usuelles : si l'on range les valeurs possibles de Δ dans l'ordre s, p, d, f, g, etc., les valeurs de Δ figurant dans une formule de série sont toujours des valeurs immédiatement voisines. Si, au lieu de désigner les valeurs de Δ par s, p, d..., on les désignait par 1, 2, 3..., on pourrait dire : « quand on passe d'un terme à l'autre dans une formule de série, Δ varie d'une unité en plus ou en moins ». C'est là ce qu'on nomme « une règle de sélection » montrant que, du moins dans les conditions usuelles d'émission, un grand nombre de combinaisons entre termes spectraux ne correspondent pas à des raies réellement observables.

2. ***Le spectre de doublets des alcalins.*** — Le tableau (2) des termes spectraux n'est qu'une toute première approximation : la réalité est malheureusement loin d'être aussi simple. Les progrès de la spectroscopie ont montré en particulier que les raies considérées comme simples dans le schéma précédent de Rydberg-Ritz étaient en réalité formées par un groupe de raies voisines constituant des doublets ou des triplets ou plus généralement des multiplets. Comme nous ne pouvons faire ici un exposé complet de cette question vraiment compliquée, nous nous bornerons à examiner le spectre des éléments alcalins où les raies forment des doublets.

L'étude du spectre des alcalins a montré que pour ces éléments la plupart des termes spectraux du schéma de Rydberg-Ritz sont dédoublés. D'une façon plus précise, la valeur s de la quantité Δ reste toujours unique, mais les valeurs p, d, f, g... sont toutes doubles : il y a deux valeurs p_1 et p_2 très

rapprochées, deux valeurs d_1 et d_2 très voisines etc. Il y a donc lieu de remplacer le tableau (2) par le tableau suivant :

$(1, s)$	$(2, s)$	$(3, s)$	$(4, s)$	$(5, s)$	
	$(2, p_1)$	$(3, p_1)$	$(4, p_1)$	$(5, p_1)$	
	$(2, p_2)$	$(3, p_2)$	$(4, p_2)$	$(5, p_2)$	
		$(3, d_1)$	$(4, d_1)$	$(5, d_1)$	
		$(3, d_2)$	$(4, d_2)$	$(5, d_2)$	(4)
			$(4, f_1)$	$(5, f_1)$	
			$(4, f_2)$	$(5, f_2)$	
				$(5, g_1)$	
				$(5, g_2)$	

Les séries optiques usuelles s'obtiennent en combinant des termes spectraux où Δ prend deux valeurs voisines sur la liste : $s, p, d, f, g \dots$ C'est la règle de sélection déjà signalée. Mais ici nous en trouvons une autre supplémentaire. Par exemple, considérons la série diffuse de formule

$$\nu = (2, p) - (n, d) \tag{5}$$

et envisageons les raies de cette série pour lesquelles $n = 3$. *A priori*, ce groupe de raies pourrait comprendre les 4 raies suivantes :

$$(2, p_1) - (3, d_1); (2, p_1) - (3, d_2); (2, p_2) - (3, d_1); (2, p_2) - (3, d_2). \tag{6}$$

Or l'expérience prouve que la seconde de ces raies ne se manifeste jamais dans les conditions usuelles. C'est là la manifestation d'une nouvelle règle de sélection dont nous donnerons l'énoncé précis après avoir introduit la notion de nombres quantiques dans la définition des termes spectraux du tableau (4).

3. ***Interprétation du schéma de Rydberg-Ritz par la théorie de Bohr.*** — La théorie de Bohr ayant donné d'excellents résultats pour l'hydrogène et l'hélium ionisé, il était tout naturel de chercher à l'étendre à des atomes plus complexes et naturellement le premier but à atteindre dans cette tentative devait être l'interprétation du tableau (2) des spectres spectraux de Rydberg-Ritz.

Lorsqu'on cherche à étendre la théorie de Bohr à des atomes contenant plus d'un électron-planète, on se heurte à de graves difficultés. Le problème dynamique devient compliqué, l'application des règles de quanta devient incertaine. Cependant l'analogie générale des spectres de tous les éléments et l'intervention dans tous de la constante R de Rydberg a conduit à penser que le schéma planétaire si utile pour l'atome d'hydrogène devait pouvoir être utilisé au moins dans une certaine mesure pour tous les autres éléments. Pour le faire, on a commencé par admettre une hypothèse très grossière : dans l'atome de rang N, on a considéré $N-1$ des électrons-planètes comme tournant au voisinage du noyau et formant autour de lui une « carcasse électronique » tandis que le N^e électron-planète dit « électron optique » aurait son orbite à l'extérieur de la carcasse électronique. Les passages de l'électron optique d'un état stationnaire à un autre détermineraient le spectre optique de l'élément. Grâce à l'hypothèse de la carcasse électronique, on peut admettre approximativement que l'action du noyau de charge $+Ne$ et de la carcasse de charge $-(N-1)e$ sur l'électron optique est équivalente par compensation à l'action d'un centre de charge $+e$. On est ainsi ramené au problème de l'atome d'hydrogène avec un seul nombre quantique. En deuxième approximation, on cherchera avec Sommerfeld à tenir compte de ce que la charge du noyau et celle de la carcasse ne se compensent pas exactement à une unité près ; la trajectoire de l'électron ne sera plus fermée, il faudra introduire un second nombre quantique k et l'on pourra ainsi rendre compte du passage des termes ayant la forme de Balmer aux termes de la forme (n, Δ). Sommerfeld a calculé, assez grossièrement il est vrai, les termes (n, Δ) que l'on obtient de cette façon et retrouvé les formules de Rydberg et de Ritz. La tentative théorique de Sommerfeld et d'autres plus élaborées que l'on trouvera dans les livres sur l'ancienne théorie des quanta [1] ont conduit à faire correspondre les diverses

[1] Voir en particulier Léon Brillouin : *L'atome de Bohr*, Paris, Presses universitaires (1931), ch. XII.

valeurs de Δ (s, p, d...) aux valeurs du nombre quantique azimutal k de Sommerfeld de la façon suivante :

$k=$	1	2	3	4	5	6	......	
$\Delta=$	s	p	d	f	g	h	......	(7)

Ceci nous permet d'écrire tout terme spectral (n, Δ) sous la forme (n, k) : par exemple le terme (2, p) s'écrit (2, 2).

Comme le nombre azimutal k est toujours inférieur ou égal au nombre total n et ne peut prendre la valeur zéro, on s'explique pourquoi pour $n=1$ la quantité Δ ne peut avoir que la valeur s, pour $n=2$ elle ne peut avoir que les valeurs s et p, etc. Les particularités du tableau (2) se trouvent ainsi entièrement expliquées. De plus par des raisonnements basés sur le principe de correspondance de Bohr, l'ancienne théorie des quanta a pu montrer que les seules transitions qui doivent avoir lieu effectivement sont celles pour lesquelles $\delta k = \pm 1$. C'est bien la règle de sélection que nous avions signalée pour le tableau (2) puisqu'une variation d'une unité en plus ou en moins pour k correspond d'après (7) à un déplacement d'un rang dans la série des valeurs possibles de Δ.

4. *Les spectres de doublets et le nomble quantique* j. — Les théories de Bohr et de Sommerfeld, en introduisant les deux nombres quantiques n et k, sont donc parvenues à interpréter les termes spectraux (2) du schéma de Rydberg-Ritz ; mais nous avons vu que ce tableau était insuffisant parce que les termes spectraux qui y figurent comme simples sont dans la réalité multiples. Il est naturel de penser que, pour caractériser chacun des termes spectraux multiples correspondant à un même terme du tableau (2), il faut introduire un troisième nombre quantique. C'est ce que l'on a fait, déjà du temps de l'ancienne théorie des quanta, en introduisant d'une manière tout à fait empirique à côté des nombres n et k un troisième nombre j ; à ce nouveau venu Sommerfeld a donné alors le nom de « nombre quantique interne » nom qui aujourd'hui ne paraît plus bien justifié.

Sans exposer ici comment l'introduction du nombre quantique j a permis de classer les multiplets optiques complexes, nous nous bornerons à étudier à ce point de vue les doublets des alcalins. Dans ce cas, nous l'avons vu, chaque terme spectral (n, Δ) est en général double ; ceci veut dire d'après les idées exposées au dernier paragraphe qu'à une même valeur du nombre k correspondent pour Δ au lieu d'une valeur unique deux valeurs très voisines. Ces deux valeurs voisines de Δ qui correspondent à une valeur déterminée de n et de k doivent être caractérisées par des valeurs différentes du nombre quantique j. On a été amené ainsi à attribuer aux deux termes spectraux voisins les deux valeurs de j :

$$j = k - 1 \pm \tfrac{1}{2} = l \pm \tfrac{1}{2} \tag{8}$$

en posant la notation :

$$k - 1 = l \tag{9}$$

dont nous verrons plus tard l'utilité. On a admis de plus que le nombre j ne peut pas prendre de valeurs négatives de sorte que pour $k = 1$ on a seulement la valeur $j = \frac{1}{2}$; ainsi se trouve expliquée l'unicité des termes s dans le tableau (4). La formule (8) conduit immédiatement au tableau suivant de correspondance entre les nombres quantiques k et j et les valeurs de Δ :

k	1	2	3	4	5	. . .
l	0	1	2	3	4	
j	$\frac{1}{2}$	$\frac{1}{2}$ $\frac{3}{2}$	$\frac{3}{2}$ $\frac{5}{2}$	$\frac{5}{2}$ $\frac{7}{2}$	$\frac{7}{2}$ $\frac{9}{2}$	. . . (10)
Δ	s	p_1 p_2	d_1 d_2	f_1 f_2	g_1 g_2	. . .

On peut donc maintenant représenter chaque terme spectral du tableau (4) par le symbole (n, k, j) et l'on obtient, à la place de (4), le tableau suivant :

$(1, 1, \frac{1}{2})$	$(2, 1, \frac{1}{2})$	$(3, 1, \frac{1}{2})$	$(4, 1, \frac{1}{2})$	$(5, 1, \frac{1}{2})$ . . .
	$(2, 2, \frac{1}{2})$	$(3, 2, \frac{1}{2})$	$(4, 2, \frac{1}{2})$	$(5, 2, \frac{1}{2})$ . . .
	$(2, 2, \frac{3}{2})$	$(3, 2, \frac{3}{2})$	$(4, 2, \frac{3}{2})$	$(5, 2, \frac{3}{2})$ . . .
		$(3, 3, \frac{3}{2})$	$(4, 3, \frac{3}{2})$	$(5, 3, \frac{3}{2})$ . . .
		$(3, 3, \frac{5}{2})$	$(4, 3, \frac{5}{2})$	$(5, 3, \frac{5}{2})$ . . . (11)

$$\begin{array}{ll} (4,\ 4,\frac{5}{2}) & (5,\ 4,\frac{5}{2})\cdots \\ (4,\ 4,\frac{7}{2}) & (5,\ 4,\frac{7}{2})\cdots \\ & (5,\ 5,\frac{7}{2})\cdots \\ & (5,\ 5,\frac{9}{2})\cdots \end{array}$$

Nous avons remarqué que pour les termes spectraux des alcalins il existe, en dehors de la règle de sélection valable pour le tableau (2) (c'est-à-dire $\delta k = \pm 1$), une autre règle du même genre puisque par exemple la deuxième des 4 raies (6) ne se produit pas dans les conditions usuelles. Cette nouvelle règle s'exprime à l'aide du nombre j par la formule $\delta j = \begin{cases} +1 \\ 0 \\ -1 \end{cases}$. Nous avons donc au total les deux règles suivantes:

$$\delta k = \begin{cases} +1 \\ -1 \end{cases} \qquad \delta j = \begin{cases} +1 \\ 0 \\ -1 \end{cases} \tag{12}$$

Vérifions ces règles sur les 4 raies $(2p - 3d)$ données en (6). Avec nos notations actuelles elles s'écrivent :

$$\begin{array}{ll} (2,\ 2,\frac{3}{2}) - (3,\ 3,\frac{5}{2}) & (2,\ 2,\frac{1}{2}) - (3,\ 3,\frac{5}{2}); \\ (2,\ 2,\frac{3}{2}) - (3,\ 3,\frac{3}{2}) & (2,\ 2,\frac{1}{2}) - (3,\ 3,\frac{3}{2}). \end{array} \tag{13}$$

Pour toutes les quatre, on a $\delta k = +1$ mais l'une d'entre elles $(2, 2, \frac{1}{2}) - (3, 3, \frac{5}{2})$ ne satisfait pas à la deuxième règle (12). C'est bien celle qui expérimentalement ne se manifeste pas.

Dans l'ancienne théorie des quanta, le sens du nombre j est resté inconnu, les tentatives faites pour l'interpréter s'étant montrées insuffisantes. L'hypothèse de l'électron magnétique et tournant d'Uhlenbeck et Goudsmit a révélé sa vraie signification et nous le verrons apparaître tout naturellement dans la théorie de Dirac.

CHAPITRE III

Les spectres de Rayons X
et les théories de Bohr et Sommerfeld

1. *La loi de Moseley et la théorie de Bohr*

En gros les spectres de Rayons X présentent les mêmes caractères généraux que les spectres optiques. Le principe de combinaison leur est applicable ; on y rencontre des séries correspondant à la combinaison d'un terme spectral fixe avec un terme spectral variable ; on y retrouve des règles de sélection.

En première approximation, on peut représenter les termes spectraux X par l'expression :

$$\frac{R\,N^2}{n^2} \qquad n = 1, 2 \; . \; . \; . \tag{1}$$

où R est la constante de Rydberg et N le nombre atomique du corps simple émetteur. C'est là la loi de Moseley. La présence ici encore de la constante R indique l'étroite parenté des spectres Röntgen avec les spectres optiques.

En réalité, la forme (1) des termes spectraux n'est qu'une toute première approximation et les véritables termes spectraux X présentent exactement le même degré de complexité que les termes spectraux optiques des alcalins. Les doublets qu'on rencontre ici ont paru d'abord pouvoir s'expliquer par la théorie de structure fine de Sommerfeld, mais un examen

plus attentif a montré que cette théorie était insuffisante et les choses n'ont été finalement remises en ordre que par la théorie de l'électron magnétique en particulier sous la forme de Dirac.

Dans les anciennes idées simples de la théorie de Bohr, l'émission des Rayons X est liée à une réorganisation de la carcasse électronique. Bohr, nous l'avons vu au dernier chapitre, considérait l'émission des raies lumineuses comme due aux passages de l'électron le plus extérieur d'une orbite stable à une autre, l'ensemble des électrons plus intérieurs formant une sorte de carcasse par l'entrelacement de leurs orbites. L'émission des Rayons X correspondrait alors aux modifications que peut subir ce système interne en passant d'un état stable à un autre. Naturellement le calcul exact de ces états stables même dans l'ancienne théorie des quanta est inextricable.

On est parvenu à un calcul approximatif, très grossier du reste, de la manière suivante. On admet que les orbites électroniques de la carcasse sont individuellement caractérisées par un ou plusieurs nombres quantiques, ce qui revient en un certain sens à négliger l'interaction des électrons en les considérant séparément. On dit alors que les électrons possédant des énergies égales ou presque égales forment une « couche » ; une couche peut d'ailleurs comprendre des orbites de types différents caractérisées par des ensembles différents de nombres quantiques, mais les énergies de ces orbites doivent être très voisines.

Les calculs faits pour l'hydrogène nous indiquent que les orbites d'une même couche doivent être caractérisées par un même nombre quantique total n. L'expérience a révélé dans les atomes l'existence de couches que l'on a pris l'habitude de désigner par les lettres consécutives de l'alphabet K, L, M, etc. On s'est rendu compte que pour la couche K on a $n=1$, pour la couche L on a $n=2$, etc.

Ce sont les orbites de la couche K qui ont la plus petite énergie. D'après un principe bien connu, il semblerait donc qu'à l'état normal tous les électrons de l'atome devraient se

trouver dans la couche K. L'étude des spectres X et la périodicité des propriétés chimiques dans la table de Mendeleeff prouvent qu'il n'en est certainement pas ainsi. Nous devons admettre une sorte de « saturation des couches », c'est-à-dire admettre qu'il existe un nombre maximum d'électrons dans chaque couche. L'examen critique des renseignements que l'expérience fournit à ce sujet a permis d'énoncer la règle suivante :

« Le nombre maximum d'électrons qui peuvent appartenir à la couche définie par le nombre quantique total n est égal à $2n^2$. »

Nous verrons plus loin comment on a pu préciser la répartition des électrons entre les niveaux d'énergie appartenant à une même couche.

Rappelons maintenant comment MM. Bohr et Kossel se sont représentés l'émission des Rayons X. Un agent extérieur (projectile matériel ou radiation) arrivant sur un atome dans l'état normal peut arracher et projeter au dehors un des électrons de la carcasse. L'atome reste alors dans un état anormal; il a subi une « ionisation profonde ». Soient alors W_0 l'énergie normale minimum de l'atome, W_1 son énergie après l'ionisation profonde.

La différence $W_1 - W_0$ a été fournie par l'agent ionisant : nous verrons qu'elle correspond (au facteur $\frac{1}{h}$ près) à la fréquence limite d'une série spectrale X. L'atome présentant alors une place libre dans sa carcasse, une réorganisation de celle-ci va pouvoir se produire spontanément car un électron va pouvoir abandonner la place qu'il y occupait au début et venir occuper la place libre. Naturellement cette transformation ne pourra s'accomplir que si elle correspond à une diminution de l'énergie totale de l'atome et devra s'accompagner d'un rayonnement qui appartiendra au domaine X.

Bref l'origine des rayons X dans cette conception serait le passage d'un électron d'une certaine couche à une autre de moindre énergie où une ionisation préalable a créé une place vide.

Après la réorganisation de sa carcasse, l'atome se trouve dans un état d'ionisation moins profonde et possède une énergie W_2 intermédiaire entre W_0 et W_1.

Envisageons ce second état ionisé et souvenons-nous que tous les électrons sont considérés comme semblables : ce nouvel état sera identique à celui qu'aurait créé l'agent extérieur si, au lieu d'expulser le premier électron, il avait expulsé le second, celui qui s'est déplacé dans la réorganisation. Donc l'énergie $W_2 - W_0$ représente le travail correspondant à l'ionisation moins profonde. La règle de Bohr conduit ainsi à attribuer au rayonnement X émis lors de la réorganisation la fréquence :

$$\nu = \frac{1}{h}(W_1 - W_2) = \frac{1}{h}\left[(W_1 - W_0) - (W_2 - W_0)\right] \qquad (2)$$

Dressons la liste des énergies W_1... de l'atome dans ses divers états d'ionisation profonde en en retranchant l'énergie W_0 de l'état normal : nous obtenons ce qu'on peut nommer les niveaux d'énergie de l'atome comptés à partir de l'état normal. En divisant par h, on obtient les termes spectraux X :

$$\frac{W_1 - W_0}{h} \qquad \frac{W_2 - W_0}{h} \; \ldots \; \frac{W_n - W_0}{h} \; \ldots \qquad (3)$$

Les termes spectraux X sont donc égaux aux travaux d'ionisation divisés par h. Les fréquences des raies X sont de la forme :

$$\nu = \left[\frac{W_n - W_0}{h} - \frac{W_m - W_0}{h}\right] \qquad (4)$$

Les raies d'une même série correspondent à un même W_n c'est-à-dire à un même niveau d'arrivée pour l'électron lors de la réorganisation. Le second terme de (4) diffère pour les différentes raies d'une même série et tend vers zéro pour les raies d'ordre élevé : la fréquence des raies quand l'ordre s'élève dans la série tend vers une limite égale au terme spectral $\frac{W_n - W_0}{h}$ qui caractérise la série.

La loi de Moseley nous apprend que dans la suite des éléments chacun des termes spectraux (3) varie en première et assez grossière approximation comme le carré du nombre atomique. Pour retrouver cette loi avec la théorie de Bohr, on admet qu'un électron de la carcasse peut être regardé comme soumis à un champ central égal à $(N-z)\frac{e}{r^2}$, le terme $\frac{ze}{r^2}$ représentant très grossièrement l'action répulsive des électrons plus rapprochés du noyau. On est ainsi ramené une fois de plus au cas hydrogénoïde et en posant $N' = N - z$, on obtient pour les termes spectraux la forme générale :

$$R\frac{N'^2}{n^2} \qquad (5)$$

Pour les couches profondes des atomes lourdes, z est petit et l'on peut confondre N avec N' ; on retrouve alors la forme (1) de la loi de Moseley. Il est bien évident que le calcul est peu rigoureux : la théorie de Bohr indique plutôt qu'elle ne démontre la loi de Moseley.

2. ***Analyse sommaire des spectres* X.** — Le premier fait qui a frappé les spectroscopistes dans le domaine Röntgen est le suivant : les raies X se divisent en groupes très nettement isolés les uns des autres dans l'échelle des fréquences ou des longueurs d'onde. On a d'abord considéré chacun de ces groupes de raies comme formant une seule série et l'on a distingué ces séries par les lettres consécutives de l'alphabet à partir de la lettre K : série K, série L, série M, etc. L'ensemble des séries se déplace vers les hautes fréquences quand on passe d'un élément à un autre plus lourd et à peu près comme le carré du nombre atomique N (Moseley). Un examen plus attentif a ensuite montré que les groupes de raies appelés d'abord série L, série M, etc. se décomposent en réalité en sous-groupes formant les véritables séries dont les raies d'ailleurs s'enchevêtrent sur les spectrogrammes. Il

existe donc plusieurs niveaux L, plusieurs niveaux M, etc. d'énergies légèrement différentes. Le niveau K, lui, est unique.

L'expérience a indiqué qu'il existe 3 niveaux L, 5 niveaux M, 7 niveaux N... A chacun de ces niveaux correspond bien entendu un terme spectral. On distingue les niveaux et les termes spectraux associés par des chiffres romains, les chiffres les plus bas correspondant aux niveaux les plus profonds, c'est-à-dire aux travaux d'ionisation les plus élevés.

Voici à titre d'exemple le tableau des raies X des séries K L M pour les éléments lourds :

Termes	Série K	Séries L			Séries M				
	K	L_I	L_{II}	L_{III}	M_I	M_{II}	M_{III}	M_{IV}	M_V
L_I	.	.	.	.	.	.	.	.	.
L_{II}	α_2	.	.	.	.	.	.	.	.
L_{III}	α_1	.	.	.	.	.	.	.	.
M_I	.	.	η	l	.	.	.	.	.
M_{II}	β_2	β_4	.	.	.	.	.	.	.
M_{III}	β_1	β_3	.	.	.	.	.	.	.
M_{IV}	.	.	β_1	α_2	.	.	.	.	.
M_V	.	.	.	α_1	.	.	.	.	.
N_I	.	.	γ_5	β_6	.	.	.	.	.
N_{II}	γ_2	γ_2	.	.	.	.	.	.	.
N_{III}	γ_1	γ_3	.	.	.	.	.	.	.
N_{IV}	.	.	γ_1	β_2'	.	ε	.	.	.
N_V	.	.	.	β_2	.	.	γ	.	.
N_{VI}	.	.	.	.	.	.	.	β	α_2
N_{VII}	.	.	.	.	.	.	.	.	α_1
O_I	.	.	γ_{11}	β_1	.	.	.	.	.
O_{II}	.	γ_4'	.	.	.	.	.	.	.
O_{III}	δ_1	γ_4	.	.	.	.	.	.	.
O_{IV}	.	.	γ_2	.	.	.	.	.	.
O_V	.	.	.	β_5	.	.	δ	.	.
P_I	.	.	γ_2'	β_5'	.	.	.	.	.
P_{II}	.	.	.	.	.	.	.	.	.
P_{III}	.	γ_8	.	.	.	.	.	.	.

Il est facile d'interpréter ce tableau. Le nom de chaque raie est inscrit au croisement d'une ligne et d'une colonne : sa fréquence est la différence du terme spectral situé en tête de la colonne et du terme spectral placé en tête de la ligne. Les raies d'une même série se trouvent ainsi placées dans une même colonne.

La série K unique est formée d'une suite de doublets : $\alpha_1 - \alpha_2$, $\beta_1 - \beta_2$, $\gamma_1 - \gamma_2$... dont l'intervalle va en resserrant. Tous ces doublets se dilatent quand le nombre atomique croît et leur écartement en fréquence augmente comme la 4e puissance du nombre atomique. On appelle cette sorte de doublets « doublets réguliers » ou « doublets de Sommerfeld ». La série L_I a une structure analogue à celle de la série K ; elle aussi est formée de doublets réguliers dont l'écartement diminue quand on s'élève dans la série.

Au contraire les séries L_{II} et L_{III} présentent un aspect très différent des précédentes. Les raies homologues de ces deux séries (figurant sur une même ligne du tableau) forment des doublets à écartement constant d'un bout à l'autre de la série ($\eta - l$, $\beta_1 - \alpha_2$, $\gamma_5 - \beta_6$, $\gamma_1 - \beta_2'$, $\gamma_{11} - \beta_7$, $\gamma_2' - \beta_3'$). Cet écartement constant est égal à $\nu_{L_{II}} - \nu_{L_{III}}$ et varie comme N^4 dans la série des éléments. Ces doublets sont encore des « doublets réguliers »

On peut résumer d'une façon générale les lois des spectres X en disant : Les termes spectraux X varient en gros comme N^2 (loi de Moseley) ; la différence de deux termes spectraux consécutifs de même nom (L, M...) dont le premier a un indice pair, par exemple $\nu_{L_{II}} - \nu_{L_{III}}$, varie comme N^4 et donne naissance aux doublets réguliers.

Nous compléterons encore ceci en faisant la remarque suivante : la différence de deux termes spectraux de même nom dont le premier a un indice impair, par exemple $\nu_{L_I} - \nu_{L_{II}}$, varie dans la suite des éléments de telle façon que $\sqrt{\nu_{L_I}} - \sqrt{\nu_{L_{II}}}$ reste constant et définit ce que l'on nomme un « doublet irrégulier ».

3. ***Classification des termes spectraux* X.** — Le schéma de Bohr et Kossel conduit à admettre que l'émission d'une série spectrale X est consécutive à une ionisation profonde de l'atome. Si cette ionisation fait passer l'énergie de l'atome de sa valeur normale W_0 à la valeur W_n, le terme spectral caractéristique de la série est $\frac{W_n - W_o}{h}$. Admettons que chaque orbite électronique puisse être quantifiée séparément ; on peut alors présumer que la quantification introduira trois nombres quantiques, n, k, j comme pour les spectres optiques. L'ionisation profonde à laquelle correspond l'énergie W_n peut alors être repérée par le symbole (n, k, j) formé avec les trois nombres quantiques qui définissent l'orbite parcourue, avant l'ionisation, par l'électron expulsé. De là résulte que nous devons pouvoir faire correspondre chaque terme spectral X à l'un des symboles (n, k, j) déjà rencontrés dans les spectres optiques. Or l'étude des séries X montre qu'elles ont tout à fait la même structure que les spectres des alcalins de sorte qu'il doit exister une correspondance univoque entre les termes spectraux X et les termes spectraux des alcalins (tableau (II) du chapitre précédent). Voici cette correspondance :

		$\left(1,1,\frac{1}{2}\right)=K$ $1\,s$			
$\left(2,1,\frac{1}{2}\right)=L_{I}$ $2\,s$	$\left(3,1,\frac{1}{2}\right)=M_{I}$ $3\,s$	$\left(4,1,\frac{1}{2}\right)=N_{I}$ $4\,s$	$\left(5,1,\frac{1}{2}\right)=O_{I}$ $5\,s$	$\left(6,1,\frac{1}{2}\right)=P_{I}$ $6\,s$	
$\left(2,2,\frac{1}{2}\right)=L_{II}$ $2\,p_1$	$\left(3,2,\frac{1}{2}\right)=M_{II}$ $3\,p_1$	$\left(4,2,\frac{1}{2}\right)=N_{II}$ $4\,p_1$	$\left(5,2,\frac{1}{2}\right)=O_{II}$ $5\,p_1$	$\left(6,2,\frac{1}{2}\right)=P_{II}$ $6\,p_1$	
$\left(2,2,\frac{3}{2}\right)=L_{III}$ $2\,p_2$	$\left(3,2,\frac{3}{2}\right)=M_{III}$ $3\,p_2$	$\left(4,2,\frac{3}{2}\right)=N_{III}$ $4\,p_2$	$\left(5,2,\frac{3}{2}\right)=O_{III}$ $5\,p_2$	$\left(6,2,\frac{3}{2}\right)=P_{III}$ $6\,p_2$	
	$\left(3,3,\frac{3}{2}\right)=M_{IV}$ $3\,d_1$	$\left(4,3,\frac{3}{2}\right)=N_{IV}$ $4\,d_1$	$\left(5,3,\frac{3}{2}\right)=O_{IV}$ $5\,d_1$		(6)
	$\left(3,3,\frac{5}{2}\right)=M_{V}$ $3\,d_2$	$\left(4,3,\frac{5}{2}\right)=N_{V}$ $4\,d_2$	$\left(5,3,\frac{5}{2}\right)=O_{V}$ $5\,d_2$		
		$\left(4,4,\frac{5}{2}\right)=N_{VI}$ $4\,f_1$			
		$\left(4,4,\frac{7}{2}\right)=N_{VII}$ $4\,f_2$			

On remarquera sur ce tableau que l'édification des couches O et P n'est pas achevée même pour les atomes lourds faute d'électrons.

Guidés par l'analogie optique, nous voyons maintenant que la série K a pour formule spectrale $(1, s)-(n, p)$: c'est donc une première série principale formée comme celle des alcalins de doublets qui vont en se resserrant. La série L est une « seconde série principale » de formule $(2, s)-(n, p)$ et de structure analogue. L'analogie des séries K et L_I se trouve ainsi interprétée.

A notre point de vue actuel, les séries L_{II} et L_{III} se décomposent en deux ensembles de raies. Le premier ensemble est formé par les doublets réguliers à écartement constant et à composantes simples $\eta - l$, $\gamma_5 - \beta_6$, $\gamma_{11} - \beta_7$, $\gamma_2' - \beta_5'$ et est l'analogue de la série étroite $(2, p)-(n, s)$ des alcalins. Le second ensemble de raies forme une série diffuse de formule $(2, p)-(n, d)$ et renferme les raies les plus intenses du groupe L ; elles forment des doublets à écartement constant mais qui présentent une structure fine due à la complexité des termes d. On peut d'ailleurs sur le tableau des raies X donné plus haut vérifier comme pour les spectres optiques les règles de sélection $\delta k = \pm 1, \delta j = \begin{cases} 0 \\ \pm 1 \end{cases}$. Le classement des raies X par analogie avec les spectres de doublets des alcalins, classement qui s'étend aux séries M et N, fournit un schéma parfaitement clair et cohérent et paraît imposer la répartition des nombres quantiques entre niveaux donnée par le tableau (6).

4. *Interprétation théorique. Formule de structure fine de Sommerfeld.* — Voyons maintenant comment on a cherché à rendre compte de l'allure des termes spectraux X dans l'ancienne théorie des quanta. L'image la plus ancienne et la plus simpliste a consisté à regarder l'atome de nombre atomique N comme contenant N électrons-planètes tournant sur des orbites circulaires coplanaires et concentriques (cercle K, cercle L, etc.). Pour tenir compte grossièrement de l'action

mutuelle des électrons, on supposait qu'elle pouvait se traduire par un simple effet d'écran ayant pour conséquence une diminution apparente de la charge du noyau. Ainsi un électron K serait soumis à une force $(N - k)\frac{e^2}{r^2}$, un électron L à une force $(N - l)\frac{e^2}{r^2}$ etc., k, l... étant appelés « nombres d'écran ». Les calculs valables pour l'atome d'hydrogène peuvent alors s'appliquer ici sans difficultés et l'on trouve les termes spectraux :

$$\frac{-Rh(N-k)^2}{1^2} \qquad \frac{-Rh(N-l)^2}{2^2} \dots \tag{7}$$

c'est-à-dire à peu près la loi de Moseley. Cette première théorie approximative est manifestement insuffisante pour beaucoup de raisons et en particulier parce qu'elle prévoit seulement un niveau L, un niveau M, etc.

Pour expliquer la multiplicité des niveaux par couches, Sommerfeld a introduit ici encore la dynamique de Relativité en considérant toutes les orbites circulaires ou elliptiques. Nous avons vu que par ce procédé il avait trouvé pour l'atome de rang N ionisé (N — 1) fois les termes spectraux :

$$\frac{E_{nk}}{h} = -\frac{RN^2}{n^2}\left[1 + \frac{\alpha^2 N^2}{n^2}\left(\frac{n}{k} - \frac{3}{4}\right)\right] \tag{8}$$

où α est la constante de structure fine et k le nombre quantique azimutal tel quel $0 < k \leqslant n$. Si l'on admet qu'on peut grossièrement représenter la répulsion mutuelle des électrons par un nombre d'écran, le terme spectral X caractérisé par les nombres quantiques n et k est donné par :

$$\frac{E_{nk}}{h} = -\frac{R(N - z_{nk})^2}{n^2}\left[1 + \alpha^2\frac{(N - z_{nk})^2}{n^2}\left(\frac{n}{k} - \frac{3}{4}\right)\right] \tag{9}$$

z_{nk} étant le nombre d'écran relatif à la trajectoire caractérisée par n et k. Comme pour une valeur donnée de n le nombre k peut prendre n valeurs distinctes, la formule (9) fait prévoir 1 niveau K, 2 niveaux L, 3 niveaux M, 4 niveaux N, etc. ce qui est, nous l'avons vu au paragraphe précédent,

encore insuffisant. La théorie de Sommerfeld, plus complète que la théorie primitive, est donc encore trop étroite parce qu'elle n'introduit pas le nombre quantique j.

Néanmoins, malgré son insuffisance évidente, la théorie de Sommerfeld a paru remporter un très grand succès par l'explication quantitative des doublets réguliers. Prenons pour exemples les doublets réguliers des séries L. Ils proviennent de la combinaison d'un même terme M, N, etc. avec les termes L_{II} et L_{III} respectivement.

Les deux raies du doublet ont donc des fréquences de la forme $\nu_{L_{II}} - \nu_i$ et $\nu_{L_{III}} - \nu_i$; leur différence de fréquence $\delta\nu_L$ ou écartement du doublet est $\nu_{L_{II}} - \nu_{L_{III}}$. Dans sa théorie, Sommerfeld laissait de côté comme provisoirement inexpliqué le niveau L_I et il attribuait aux niveaux L_{II} et L_{III} les nombres quantiques $n=2$, $k=1$ et $n=2$, $k=2$ respectivement. La formule (9) lui donnait donc [1] :

$$\delta\nu_L = \nu_{L_{II}} - \nu_{L_{III}} = R\,\alpha^2 \frac{(N - z_L)^4}{2^4}\left(\frac{2}{1} - \frac{2}{2}\right) = \frac{\pi^4 m e^8}{2 c^2 h^5}(N - z_L)^4 \tag{10}$$

L'écartement des doublets L doit donc varier comme $(N - z_L)^4$ et c'est bien sensiblement la loi expérimentale car pour les atomes pas trop légers z_L doit être petit devant N. Si l'on pose $N=1$, $z_L=0$, on retombe sur l'écartement $\Delta\nu_{II}$ des doublets de la série de Balmer [formule (43) du chapitre 1]. On a donc :

$$\delta\nu_L = \delta\nu_H (N - z_L)^4 \tag{11}$$

La formule (11) est bien vérifiée numériquement par l'expérience à condition de poser $z_L = 3{,}5$, ce qui est une hypothèse raisonnable. Les doublets des séries M et N sont aussi bien prévus numériquement par application de la formule (9). Cette prévision exacte des doublets réguliers a paru tout d'abord un très grand succès de la théorie de structure fine de Sommerfeld, mais ensuite est apparue une grave objection que nous allons exposer.

---

(1) En confondant z_{21} et z_{22} et les posant tous deux égaux à z

5. ***Insuffisance de la théorie de la structure fine de Sommerfeld.*** — La théorie de Sommerfeld en introduisant seulement les deux nombres quantiques n et k ne nous donne pas assez de niveaux pour les spectres X. Il nous faut introduire le troisième nombre j et l'analogie optique nous contraint à répartir les nombres quantiques entre les niveaux de la manière indiquée par le tableau (6). Mais alors se présente une difficulté insurmontable pour la théorie de Sommerfeld. D'après le tableau (6), les niveaux L_{II} et L_{III} ont en effet les symboles $(2, 2, \frac{1}{2})$ et $(2, 2, \frac{3}{2})$; ils ont donc le même nombre k égal à 2 et diffèrent par leur nombre j. Mais ceci ruine l'interprétation des doublets réguliers L par la formule (9) telle que Sommerfeld l'a donnée car celle-ci suppose essentiellement que les nombres k de L_{II} et de L_{III} diffèrent d'une unité. On a donc pu croire, ce qui était d'ailleurs assez surprenant, que le succès de la théorie de Sommerfeld était purement fortuit. Le développement des théories plus récentes et en particulier de celle qui va faire l'objet de ce livre a montré depuis que ce caractère fortuit n'était qu'apparent. C'est bien l'introduction de la Relativité qui permet l'interprétation correcte des doublets réguliers, mais à condition d'introduire en même temps le caractère magnétique de l'électron et c'est l'absence de ce dernier élément dans la théorie de Sommerfeld qui causait sa faiblesse.

Nous venons de le voir, l'étude attentive des spectres X montre l'insuffisance de la théorie de structure fine de Sommerfeld. Mais même dans le cas simple de l'hydrogène, l'examen approfondi des doublets de la série de Balmer a montré que ces doublets n'étaient pas correctement interprétés par cette théorie. Considérons la raie H_α de la série de Balmer; elle provient du passage d'un état stationnaire pour lequel $n=3$ à un état stationnaire pour lequel $n=2$. Cette raie est en réalité multiple et, si nous utilisons un mode de représentation des niveaux bien connu aujourd'hui, la théorie de Sommerfeld nous donne pour la structure fine de cette raie le schéma suivant :

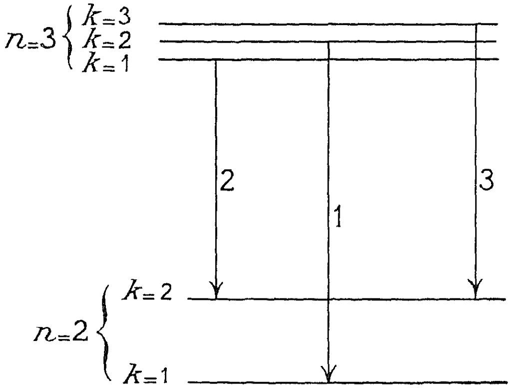


Fig. 2

En appliquant la règle de sélection $\delta k = \pm 1$,, on prévoit les trois raies indiquées dans la structure fine de H_α. Le doublet de Sommerfeld est formé par la raie 1 et l'ensemble des 2 raies de fréquences très voisines 2 et 3 généralement indiscernables.

Avec le schéma à trois nombres quantiques n, k, j, on prévoit la structure fine suivante :

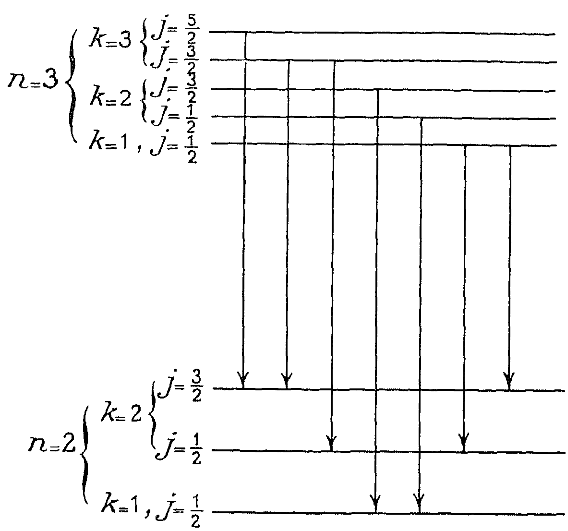


Fig. 3

Les règles de sélection $\delta k = \pm 1$ et $\delta j = 0, \pm 1$ autorisent les sept raies marquées sur la figure. Mais pour des raisons qui apparaîtront plus tard, on doit *pour l'hydrogène* avec les théories nouvelles de l'électron magnétique considérer comme confondus les niveaux ayant le même j et des k différents. On a donc le schéma simplifié suivant :

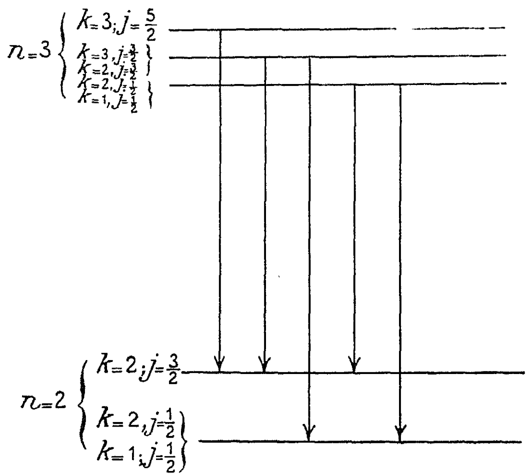


Fig. 4

Il doit donc y avoir 5 composantes dans la structure fine de H_α tandis que Sommerfeld n'en prévoyait que 3. Or l'étude soignée de cette structure fine a permis d'affirmer qu'elle comporte plus de 3 raies confirmant ainsi le nouveau schéma et non celui de Sommerfeld. L'étude des doublets de He^+ a permis de confirmer cette conclusion.

Ainsi même pour l'hydrogène et l'hélium ionisé, la théorie de Sommerfeld en dernière analyse a échoué, du moins sous sa forme primitive.

6. *Répartition des électrons entre les niveaux; règle de Stoner.* — Dans un atome lourd, les divers électrons se répartissent, nous l'avons dit, en un certain nombre de couches K, L, etc. Ces couches se décomposent elles-mêmes en sous-

couches ou niveaux : ainsi la couche L comprend les niveaux L , L_{II} et L_{III} .

Nous avons vu que chaque couche ne peut pas posséder plus d'un certain nombre d'électrons ; nous avons parlé d'une saturation des couches. Il est bien naturel de penser qu'il existe une saturation des niveaux. En s'aidant des propriétés chimiques des éléments et des caractères de leurs spectres optiques et Röntgen, divers auteurs (Bohr, Main Smith, Dauvillier et L. de Broglie, etc.) sont parvenus à étudier la répartition des électrons entre niveaux pour les divers éléments. On est ainsi arrivé à présumer quel était le nombre maximum d'électrons que chaque niveau peut avoir. M. Stoner a énoncer à ce sujet une règle qui est aujourd'hui généralement adoptée: « *Le niveau auquel correspond le symbole* (n, k, j) *contient au maximum* $2j+1$ *électrons.* »

La règle de Stoner permet de calculer le nombre maximum des électrons appartenant à la couche définie par une valeur donnée du nombre quantique total n.

En effet, comme, pour une valeur donnée de k, j prend les valeurs $k-1\pm\frac{1}{2}$, les niveaux

$$\left(n,\ k,\ k-\frac{1}{2}\right) \quad \text{et} \quad \left((n,\ k,\ k-\frac{3}{2}\right)$$

possèdent au maximum

$$2\left(k-\frac{1}{2}\right)+1+2\left(k-\frac{3}{2}\right)+1=2(2k-1)$$

électrons et l'ensemble de la couche n en possède au maximum :

$$2\sum_{1}^{n}{}_{k}(2k-1)=2[n(n+1)-n]=2n^2 \qquad (12)$$

C'est bien la loi déjà signalée au paragraphe 1.

CHAPITRE IV

Les anomalies magnétiques et l'hypothèse de l'électron tournant

1. *Les anomalies gyromagnétiques*

DES considérations électromagnétiques simples permettent d'établir une relation générale entre le moment magnétique $\overrightarrow{\mathfrak{M}}$ produit par le déplacement d'une charge sous l'action d'une force centrale et le moment de rotation constant $\overrightarrow{M}$ correspondant à ce mouvement.

Considérons une trajectoire plane fermée parcourue par un corpuscule de masse m_0 et de charge électrique ε sous l'action d'une force centrale [1]. D'après le théorème des aires, le moment de rotation $M = m_0\, v\, r \sin\alpha$ est constant.

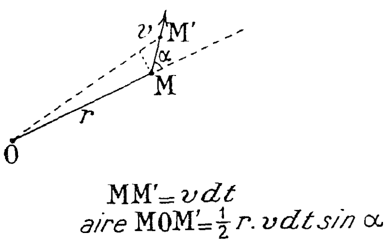


FIG. 5

Si $d\mathfrak{A}$ est l'aire balayée par le rayon vecteur pendant le temps dt, on a :

(1) Nous désignons la masse du corpuscule par m_0 au lieu de m pour éviter plus loin la confusion de cette masse avec le nombre quantique m et pour préparer le passage aux équations relativistes.

$$d\mathcal{A} = \frac{1}{2} r \cdot v\,dt \sin\alpha = \frac{1}{2m_0} \mathrm{M}\,dt \qquad (1)$$

Puisque M est une constante d'après le théorème des aires, on trouve en intégrant pour une période T du mouvement :

$$\mathcal{A} = \frac{\mathrm{M}}{2\,m_0}\,\mathrm{T} \qquad (2)$$

où $\mathcal{A}$ est l'aire totale entourée par la trajectoire.

D'autre part le mouvement de la charge ε est équivalent à l'existence d'un courant i. Ce courant est par définition égal à la quantité d'électricité qui passe par seconde à travers une aire unité normale à la trajectoire. Comme il passe la charge ε pendant le temps T, on a :

$$i = \frac{\varepsilon}{\mathrm{T}} \qquad (3)$$

en supposant ε exprimée en unités électromagnétiques. De (2) et de (3), on tire :

$$i = \frac{\mathrm{M}\,\varepsilon}{2m_0\mathcal{A}} \qquad (4)$$

Au point de vue magnétique, ce courant est équivalent à un feuillet de puissance i et de surface $\mathcal{A}$. Le moment magnétique $\mathfrak{M}$ de ce feuillet est :

$$\mathfrak{M} = i\mathcal{A} = \frac{\varepsilon\mathrm{M}}{2\,m_0} \qquad (5)$$

Cette relation est valable en grandeur et direction pour les vecteurs $\overrightarrow{\mathfrak{M}}$ et $\overrightarrow{\mathrm{M}}$; d'où :

$$\frac{\overrightarrow{\mathfrak{M}}}{\overrightarrow{\mathrm{M}}} = \frac{\varepsilon}{2\,m_0} \qquad (6)$$

On peut démontrer d'une façon plus générale que si l'on considère un ensemble de corpuscules de même masse m et de même charge ε formant un système dans un état stationnaire, la relation (6) est encore valable entre le moment

magnétique total $\vec{\mathfrak{M}}$ dû au mouvement de ces charges et le moment de rotation total $\vec{M}$ constant du système [1].

D'après la théorie de Bohr, les atomes sont des ensembles d'électrons en mouvement stationnaire. On peut donc appliquer aux atomes les relation (6) à condition de poser $\varepsilon = -\frac{e}{c}$ où e est la charge positive unité exprimée comme d'habitude en unités électro-statiques. On a donc pour les atomes la formule :

$$\frac{\vec{\mathfrak{M}}}{\vec{M}} = -\frac{e}{2m_0c} \qquad (m_0 \text{ masse de l'électron}) \qquad (7)(^2)$$

Cette formule fondamentale conduit à l'idée du « magnéton de Bohr ». En effet l'ancienne théorie des quanta pose toujours le moment total de quantité de mouvement pour un atome égal à un multiple entier de $\frac{h}{2\pi}$ d'où :

$$\mathfrak{M} = n\frac{h}{2\pi}\cdot\frac{e}{2m_0c} = n\frac{eh}{4\pi m_0 c} \qquad (n \text{ entier positif ou négatif}) \qquad (8)$$

Le moment magnétique de l'atome serait donc toujours un multiple entier d'une certaine unité dite « magnéton de Bohr » et égale à :

$$B = \frac{eh}{4\pi m_0 c} \qquad (9)$$

L'expérience bien connue de Stern et Gerlach a mis en évidence l'existence réelle du magnéton de Bohr.

---

(1) La démonstration est due à Einstein. Voir le rapport de M. DE HAAS dans *Atomes et Electrons* (*Rapports du Conseil de Physique Solvay de 1921*, Gauthier-Villars éd., Paris, 1923).

(2) Cette formule eest vraie *avec son signe* sous les conditions suivantes : on prend un système d'axes *directs* et l'on définit les moments de telle sorte que le moment de rotation d'un corpuscule tournant autour de l'axe des z dans le sens direct (inverse des aiguilles d'une montre) soit dirigé dans le sens positif des oz.

La relation (7) ne s'est cependant pas vérifiée d'une façon générale et complète. On a recherché si en soumettant un barreau magnétique à un champ magnétique ce barreau se mettait en rotation : la théorie indique en effet que le barreau doit prendre un mouvement de rotation tel que le rapport de son moment magnétique à son moment de rotation soit donné par la relation (7). Le phénomène existe bien (expérience de Einstein et de Haas) mais le rapport $\left(\frac{\mathfrak{M}}{M}\right)$ a été trouvé égal à $\frac{e}{m_0 c}$ au lieu de $\frac{e}{2 m_0 c}$! Barnett a trouvé le même rapport anormal en étudiant le phénomène inverse (création d'un moment magnétique par rotation d'un barreau). Il y a là une difficulté grave qui a conduit, nous le verrons, à concevoir l'idée du magnétisme propre de l'électron.

2. *L'effet Zeeman normal.* — Une autre difficulté qui a suggéré l'hypothèse de l'électron magnétique, c'est l'existence des anomalies de l'effet Zeeman.

Rappelons d'abord brièvement la théorie classique de l'effet Zeeman normal donnée naguère par Lorentz. Considérons le mouvement d'un corpuscule de masse m_0 et de charge ε dans un champ magnétique uniforme H. Le corpuscule subit une force égale à :

$$F = \frac{\varepsilon}{c}\left[\vec{v} \cdot \vec{H}\right] \tag{10}$$

Sir. J. Larmor a démontré un très intéressant et célèbre théorème sur le mouvement qui en résulte : « Si l'on considère un système de référence qui tourne autour de la direction du champ magnétique uniforme avec la vitesse angulaire constante :

$$o = -\frac{1}{2}\,\frac{\varepsilon H}{m_0 c} \qquad (\varepsilon \text{ en u. e. s.}) \tag{11}$$

le mouvement du corpuscule dans ce système de référence est celui qu'il aurait dans un système de référence fixe en l'absence du champ magnétique, les autres forces restant les mêmes. »

Appliquons ceci à un électron intra-atomique animé d'un mouvement périodique de fréquence ν. Si l'on crée un champ magnétique uniforme, le même mouvement pourra être exécuté par l'électron dans le système de référence tournant avec la vitesse de précession de Larmor (11). La fréquence $\frac{o}{2\pi}$ correspondant à cette précession s'ajoutera ou se retranchera de la fréquence ν du mouvement de l'électron suivant l'orientation relative du champ magnétique et de l'orbite. Il en résulte qu'un corps matériel qui en l'absence de champ magnétique émet une radiation de fréquence ν, devra en présence du champ magnétique uniforme H émettre aussi les fréquences $\nu + \frac{1}{4\pi}\frac{eH}{m_0 c}$ et $\nu - \frac{1}{4\pi}\frac{eH}{m_0 c}$. En serrant de plus près la théorie, on voit aisément qu'en observant à angle droit du champ magnétique on doit voir une raie de fréquence ν vibrant dans le sens du champ et les deux raies de fréquences $\nu \pm \frac{1}{4\pi}\frac{eH}{m_0 c}$ vibrants à angle droit tandis qu'en observant dans le sens du champ on doit seulement voir les deux dernières raies avec des vibrations circulaires de sens inverse. Ceci constitue l'effet Zeeman normal qui a en effet été observé dans un certain nombre de cas et dont la découverte, il y a une trentaine d'années, a paru constituer une éclatante vérification des conceptions électroniques de Lorentz.

L'ancienne théorie des quanta n'a rien introduit d'essentiellement nouveau quant à la prévision de l'effet Zeeman. Soit un électron dans un atome à l'état stable en l'absence de champ magnétique extérieur ; désignons par W_0 l'énergie de cet électron et par $\mathfrak{M}$ le moment magnétique de son orbite en admettant que nous ayons le droit de considérer ainsi isolément un des électrons de l'atome. En présence d'un champ magnétique extérieur uniforme H, l'énergie de l'orbite électronique sera

$$W_H = W_0 - (\overrightarrow{\mathfrak{M}} \cdot \overrightarrow{H}) \tag{12}$$

ou, en vertu de (6) :

$$W_H = W_0 + \frac{e}{2 m_0 c} (\vec{M} . \vec{H}) \tag{13}$$

Nous devons pour former $(\vec{M} . \vec{H})$ considérer la composante de $\vec{M}$ le long du champ et cette composante doit être un multiple de $\frac{h}{2\pi}$ d'après l'ancienne théorie des quanta (¹). La formule (13) devient donc :

$$W_H = W_0 + m \frac{e h H}{4\pi m_0 c} \tag{14}$$

m est un nombre quantique que l'on nomme le « nombre quantique magnétique ».

Considérons alors la raie due au passage de l'électron de l'état stable d'énergie W_0 à un état stable d'énergie moindre W'_0. En l'absence de champ extérieur, cette raie aura la fréquence :

$$\nu_0 = \frac{W_0 - W_0'}{h} \tag{15}$$

En présence du champ uniforme H, elle devient en vertu de (14) :

$$\nu_H = \frac{W_H - W'_H}{h} = \nu_0 + (m - m') \frac{e H}{4\pi m_0 c} \tag{16}$$

où $(m - m')$ peut prendre toutes les valeurs entières positives ou négatives, y compris zéro. Pour retrouver l'effet Zeeman normal, l'ancienne théorie des quanta guidée par le principe de correspondance admettait la règle de sélection $\Delta m = 0, \pm 1$. Elle retombait alors exactement sur les conclusions de la théorie classique et cela parce que dans la combinaison des formules (14) et (15) la constante h s'éliminait.

3. ***Anomalies de l'effet Zeeman; facteur de Landé.*** — La théorie de l'effet Zeeman normal donnée au paragraphe pré-

(¹) Voir Léon Brillouin : *L'Atome de Bohr*, déjà cité, p. 167.

cédent ne se vérifie que dans un petit nombre de cas. La plupart des effets Zeeman sont anormaux. Nous nous contenterons de décrire l'effet Zeeman anomal des éléments alcalins car c'est le plus simple et c'est aussi le seul dont la théorie de Dirac, impuissante jusqu'ici à tenir compte de l'interaction des électrons, ait pu rendre compte.

L'effet Zeeman des alcalins est soumis aux règles générales suivantes :

a) Les raies homologues des divers éléments alcalins présentent le même effet Zeeman ;

b) Les raies d'une même série spectrale présentent la même décomposition (Règle de Preston) ;

c) Les raies déplacées par l'action du champ magnétique forment toujours par rapport à la raie primitive une figure symétrique en fréquences et polarisations comme l'effet normal ;

d) L'intervalle spectral (différence de fréquence) entre une raie déplacée et la raie primitive est toujours égal au produit de l'intervalle normal de Lorentz par une fraction simple (Règle de Runge).

Ainsi pour une certaine raie, les composantes déplacées de l'effet Zeeman sont toujours dans l'échelle des fréquences à des distances de la raie primitive données par l'expression $\pm \frac{s}{r} \frac{eH}{4\pi m_0 c}$ où r est un entier caractéristique de la raie en question (dénominateur de Runge) et où s prend un certain nombre de valeurs entières, nombre qui détermine la multiplicité de la raie décomposée.

La règle (*d*) de Runge peut être interprétée de la façon suivante. D'après le principe de combinaison, la décomposition d'une raie traduit la décomposition des termes spectraux. On est donc conduit à penser que chaque terme spectral doit avoir son dénominateur de Runge. Soient r_1 et r_2 les dénominateurs de deux termes spectraux ; la règle de Runge pourra s'énoncer en disant que sous l'effet du champ magnétique ces termes sont modifiés respectivement des quantités

$\frac{q_1}{r_1}\Delta\nu_H$ et $\frac{q_2}{r_2}\Delta\nu_H$, q_1 et q_2 étant des nombres entiers et $\Delta\nu_H$ désignant l'intervalle normal de Lorentz $\frac{e\,H}{4\pi\,m_0\,c}$. La raie due à la combinaison des termes considérés subit en effet le déplacement spectral :

$$\left(\frac{q_1}{r_1}-\frac{q_2}{r_2}\right)\Delta\nu_H=\frac{q_1\,r_2-q_2\,r_1}{r_1\,r_2}\Delta\nu_H \qquad (17)$$

et ceci peut s'écrire sous la forme $\frac{s}{r}\,\Delta\nu_H$; nous retombons bien sur l'énoncé *d* de la règle de Runge.

La manière dont les termes spectraux se décomposent en présence d'un champ magnétique a été précisée par M. Landé qui a beaucoup contribué à débrouiller les effet Zeeman anormaux. Il a énoncé les règles suivantes :

1° Chaque terme spectral de symbole $(n,\ l,\ j)$ (en employant la notation $l=k-1$) se décompose dans un champ magnétique *faible* en $2j+1$ termes caractérisés par les nombres quantiques magnétiques *demi-entiers* : $-j$, $-(j-1)\ldots$ $+(j-1)$, $+j$;

2° En prenant pour unité l'intervalle normal $\frac{e\,H}{4\pi\,m_0\,c}$ de Lorentz, les écarts en fréquence des termes décomposés par rapport au terme primitif sont donnés par le tableau suivant :

Termes s $\left[n,0,\frac{1}{2}\right]$				-1	$+1$			
Termes p_1 $\left[n,1,\frac{1}{2}\right]$				$-\frac{1}{3}$	$+\frac{1}{3}$			
Termes p_2 $\left[n,1,\frac{3}{2}\right]$		$-\frac{6}{3}$	$-\frac{2}{3}$		$+\frac{2}{3}$	$+\frac{6}{3}$		(18)
Termes d_1 $\left[n,2,\frac{3}{2}\right]$		$-\frac{6}{5}$	$-\frac{2}{5}$		$+\frac{2}{5}$	$+\frac{6}{5}$		
Termes d_2 $\left[n,2,\frac{5}{2}\right]$	$-\frac{15}{5}$	$-\frac{9}{5}$	$-\frac{3}{5}$		$+\frac{3}{5}$	$+\frac{9}{5}$	$+\frac{15}{5}$	

Ce tableau montre que les dénominateurs de Runge sont égaux à 1 pour les termes s, à 3 pour les termes p, à 5 pour les termes d, etc. D'une façon générale, le dénominateur de Runge pour le terme $(n,\ l,\ j)$ est $2l+1$;

3° Dans les transitions qui donnent naissance aux raies

observables, le nombre quantique magnétique ne peut varier que de -1, 0 ou $+1$. Pour $\Delta m = \pm 1$, la raie émise, si elle est observée à angle droit du champ, vibre rectilignement dans la direction normale au champ ; pour $\Delta m = 0$, observée dans les mêmes conditions, elle vibre parallèlement au champ.

Le nombre magnétique m ayant les valeurs données par la règle 1, Landé a écrit les écarts inscrits dans le tableau (18) sous la forme mg où g est un nombre, le facteur de Landé, qui possède les valeurs suivantes :

	l ; $j =$	$\frac{1}{2}$	$\frac{3}{2}$	$\frac{5}{2}$	$\frac{7}{2}$
s	0	2			
p	1	$\frac{2}{3}$	$\frac{4}{3}$		
d	2		$\frac{4}{5}$	$\frac{6}{5}$	
f	3			$\frac{6}{7}$	$\frac{8}{7}$

(19)

On peut résumer le tableau (19) par la formule :

$$g = \frac{j + \frac{1}{2}}{l + \frac{1}{2}} = \frac{2j + 1}{2l + 1} \tag{20}$$

Telle est la valeur du facteur de Landé pour le terme (n, l, j) *d'un alcalin.*

Le terme spectral de l'atome alcalin qui avait, en l'absence de champ, la valeur $(n, l, j)_0$ peut prendre en présence du champ uniforme H les $2j + 1$ valeurs :

$$(n, l, j)_H = (n, l, j)_0 + mg \frac{eH}{4\pi m_0 c} \tag{21}$$

g ayant la valeur (20) et m pouvant prendre l'une quelconque des valeurs demi-entières comprises entre $-j$ et $+j$ (1).

(1) La règle (*b*) de Preston s'explique en remarquant que la décomposition Zeeman est toujours indépendante du nombre total n ; or, c'est seulement ce nombre qui varie quand on passe d'une raie à l'autre dans la même série.

L'ancienne théorie des quanta est tout aussi incapable que la théorie classique d'expliquer l'intervention du facteur g. Nous verrons qu'ici la Mécanique ondulatoire n'est pas plus heureuse. C'est seulement l'introduction du magnétisme de l'électron qui explique l'origine du facteur g dont la théorie de Dirac nous permettra de prévoir la forme correctement et sans objections.

La formule (21) n'est valable que pour un champ magnétique faible. Mais que faut-il au juste entendre par là? Nous dirons qu'un champ magnétique produisant une décomposition Zeeman anomale d'un doublet alcalin est faible si le déplacement des termes spectraux que la présence de ce champ magnétique produit est petit devant l'écartement des composantes du doublet en l'absence de champ. D'après cette définition un même champ magnétique pourra se comporter comme fort ou faible suivant les cas. Quand le champ est faible au sens qui vient d'être précisé, la formule (21) est applicable. Quand il ne peut être considéré comme faible, on a un phénomène plus compliqué qui obéit à une loi énoncée par Voigt sur laquelle nous n'insisterons pas ici. Mais si le champ est très fort, c'est-à-dire si le déplacement Zeeman qu'il produit est grand devant l'écartement normal du doublet, on obtient de nouveau un phénomène simple, l'effet Paschen-Bach : on observe alors la décomposition Zeeman *normale* centrée sur le centre de gravité des deux composantes du doublet primitif.

4. *Hypothèse de l'électron magnétique et tournant.* — Si l'on considère les lois de l'effet Zeeman anomal et si on les compare à la théorie classique de Lorentz, on s'aperçoit que pour les retrouver, il faudrait pouvoir poser, à la place de la relation (7), la relation :

$$\left|\frac{\mathfrak{M}}{\mathrm{M}}\right| = g\,\frac{e}{2\,m_0\,c} \tag{22}$$

Ceci nous conduit donc à penser qu'il existe dans la matière des moments de rotation et des moments magné-

tiques qui ne sont pas reliés par la relation (7). Les anomalies de l'effet gyromagnétique signalées à la fin du paragraphe 1 suggèrent la même conclusion. On ne peut donc pas supposer que tout le magnétisme de l'atome provient de circulations d'électrons conçus comme des charges ponctuelles. L'idée peut alors se présenter à l'esprit d'attribuer à l'électron lui-même un moment magnétique propre et un moment de rotation propre qui soient liés l'un à l'autre par une relation différente de la relation (7). C'est cette idée qu'ont mise en avant d'une manière ingénieuse MM. Uhlenbeck et Goudsmit avant même le développement de la nouvelle mécanique.

Se faisant de l'électron une image du type classique MM. Uhlenbeck et Goudsmit l'ont assimilé à une sphère d'électricité en rotation autour d'un de ses diamètres qui posséderait un moment de rotation $M = \frac{1}{2}\frac{h}{2\pi}$ et un moment magnétique égal à un magnéton de Bohr $\mathfrak{M} = \frac{eh}{4\pi m_0 c}$ de sorte que l'on ait pour $\left|\frac{\mathfrak{M}}{M}\right|$ la valeur $\frac{e}{m_0 c}$ révélée par expériences de Einstein et de Haas. Le rapport $\left|\frac{\mathfrak{M}}{M}\right|$ étant ainsi pour l'électron le double du rapport normal (7), on peut parler du « double magnétisme » de l'électron. Après le succès de l'hypothèse de MM. Uhlenbeck et Goudsmit, diverses tentatives ont été faites pour obtenir un modèle classique d'électron tournant mais ces tentatives ont perdu aujourd'hui beaucoup de leur intérêt depuis que le développement de la nouvelle Mécanique est venu nous interdire de considérer l'électron comme un petit corps bien localisé dans l'espace.

L'hypothèse de l'électron magnétique et tournant a permis dès son apparition d'entrevoir la solution des difficultés que nous avons énumérées. Prenons d'abord la question des doublets réguliers de Rayons X. MM. Uhlenbeck et Goudsmit admettaient que l'axe magnétique de l'électron était toujours

normal au plan de sa trajectoire. Comme il reste deux sens possibles pour le vecteur moment de rotation propre (souvent appelé vecteur « spin »), à chaque trajectoire de nombres quantiques n et k correspondent deux possibilités ; on comprend alors la nécessité pour achever la détermination quantique de la trajectoire stable d'introduire un nouveau nombre j susceptible de prendre deux valeurs distinctes, pour n et k donnés. Le moment total de rotation étant la somme du moment de rotation de l'électron sur son orbite égal à $k \frac{h}{2\pi}$ et du spin $\pm \frac{1}{2} \frac{h}{2\pi}$ on peut écrire :

$$M_{total} = \left(k \pm \frac{1}{2}\right) \frac{h}{2\pi} \tag{23}$$

on serait tenté de poser $j = k \pm \frac{1}{2}$. Mais nous verrons bientôt que la nouvelle Mécanique conduit à remplacer k par $l = k - 1$; on conçoit donc que la vraie formule reliant j à k soit la formule déjà signalée

$$j = k - 1 \pm \frac{1}{2} = l \pm \frac{1}{2} \tag{24}$$

Le nombre j apparaît donc comme exprimant le moment total de rotation en unités $\frac{h}{2\pi}$. Avec les idées classiques le petit aimant formé par l'électron se déplace dans le champ Coulombien ou quasi-Coulombien du noyau et de la carcasse. Tout se passe alors comme si le petit aimant était soumis au champ magnétique :

$$\vec{H} = -\frac{1}{c} \left[\vec{v} \cdot \vec{h}\right] \tag{25}$$

$\vec{h}$ étant le champ Coulombien. La formule (24) donne l'action d'un champ électrostatique sur un pôle magnétique en mouvement avec la vitesse $\vec{v}$. Le champ H étant perpendiculaire à l'orbite plane décrite par le petit aimant électronique dans le champ Coulombien, l'énergie potentielle de ce petit aimant dans le champ H est :

$$U = \pm \mathfrak{M} H \tag{26}$$

où $\mathfrak{M}$ est égal à un magnéton de Bohr suivant l'hypothèse d'Uhlenbeck et Goudsmit et où l'on devra prendre le signe + ou le signe — suivant le sens du moment $\mathfrak{M}$ normal à la trajectoire par rapport au sens de description de celle-ci. Par suite de l'existence de l'énergie potentielle (25), chaque niveau (n, k) de la théorie de Sommerfeld va se décomposer en deux niveaux (n, k, j). Des calculs faits par Uhlenbeck et Goudsmit, puis repris et perfectionnés par Thomas et Frenkel, en se servant toujours de l'ancienne théorie des quanta permettent de retrouver une loi en N^4 pour les doublets provenant de termes spectraux dont le nombre j diffère d'une unité et font ainsi disparaître la difficulté rencontrée par la théorie primitive de Sommerfeld (voir paragr. 5 du dernier chapitre). Mas ces calculs prêtent à des objections et ne parviennent à de bons résultats que moyennant des hypothèses artificielles telles par exemple que la substitution du nombre quantique l au nombre quantique k, substitution non justifiée quand on emploie l'ancienne théorie des quanta. Il est aujourd'hui certain que l'on ne peut pas traiter les problèmes intra-atomiques par les méthodes de la Mécanique ancienne et que l'on doit avoir recours à celles de la Mécanique ondulatoire. Nous n'insisterons donc pas sur les calculs de la théorie primitive de l'électron magnétique (1).

L'effet Zeeman anomal et la formule de Landé ont reçu par l'hypothèse de l'électron magnétique et tournant un commencement d'explication. Bornons-nous toujours au cas des alcalins. L'électron optique extérieur d'un élément alcalin possède un moment de rotation total (moment orbital + spin) égal à $j = l \pm \frac{1}{2}$ fois le moment unité $\frac{h}{2\pi}$, en admettant un peu arbitrairement la substitution de l à k.

Quelle sera alors l'énergie potentielle de cet électron optique en présence d'un champ magnétique extérieur H? Admettons avec l'ancienne théorie des quanta que la composante

(1) Voir Léon Brillouin, *loc. cit.*, chapitre XVI.

M_H du moment total de rotation dans le sens du champ H soit de la forme $m \frac{h}{2\pi}$, m étant le nombre quantique magnétique auquel nous imposerons de prendre l'une des valeurs demi-entières de $-j$ à $+j$. Le moment magnétique total fera donc avec le champ H un angle dont le cosinus sera $\frac{m}{j}$. Si l'électron n'avait pas de double magnétique propre, son énergie dans le champ H serait :

$$W_H = W_0 - \mathfrak{M}_H \cdot H = W_0 + \frac{e}{2 m_0 c} M_H H = W_0 + m \frac{e h H}{4 \pi m_0 c} . \tag{27}$$

On retomberait sur la formule (14) et l'effet Zeeman normal. Mais en vertu du double magnétisme de l'électron, on a (en écrivant toujours l au lieu de k) :

$$\begin{aligned} \mathfrak{M}_H &= -\frac{e}{2 m_0 c} \cdot l \frac{h}{2\pi} \cdot \frac{m}{j} \mp 2 \frac{e}{2 m_0 c} \cdot \frac{1}{2} \frac{h}{2\pi} \cdot \frac{m}{j} \\ &= -\frac{e}{2 m_0 c} \cdot \frac{h}{2\pi} \cdot \frac{l \pm 1}{j} \cdot m \cdot \end{aligned} \tag{28}$$

et par suite :

$$W_H = W_0 - \mathfrak{M}_H \cdot H = W_0 + m \frac{l \pm 1}{j} \frac{e h H}{4 \pi m_0 c} \tag{29}$$

formule qui est équivalente à la relation empirique (21) à condition de poser :

$$g = \frac{l \pm 1}{j} = \frac{j \pm \frac{1}{2}}{l \pm \frac{1}{2}} = \frac{2 j \pm 1}{2 l \pm 1} \tag{30}$$

Cette formule obtenue avec des hypothèses passablement arbitraires présente un air de parenté incontestable avec la formule empirique (20) de Landé.

L'hypothèse de l'électron magnétique développée tant bien que mal dans le cadre de l'ancienne théorie des quanta don-

nait donc des résultats intéressants. Mais aujourd'hui le succès de la nouvelle Mécanique a montré que les questions relatives à l'électron doivent être posées d'une tout autre manière. Il nous faut donc maintenant résumer les conceptions fondamentales de la Mécanique ondulatoire. Nous aurons en même temps à examiner comment on a cherché sans grand succès à introduire les idées relativistes dans la forme primitive de cette nouvelle Mécanique. Cette introduction n'a pas permis de résoudre les difficultés signalées dans les pages qui précèdent. C'est la théorie de Dirac qui, en introduisant simultanément dans le cadre de la Mécanique ondulatoire le principe de Relativité et le magnétisme propre de l'électron, a pu aplanir ces difficultés.

CHAPITRE V

Résumé des principes de la mécanique ondulatoire

1. *Point de vue de la nouvelle Mécanique*

DANS l'ancienne Mécanique, on considérait les corpuscules ou points matériels comme de petits objets de dimensions négligeables ayant à chaque instant une position bien définie dans l'espace. Si un corpuscule est en mouvement, l'ensemble de ses positions successives constitue alors sa trajectoire. Les équations classiques de la Dynamique de Newton (ou les équations un peu modifiées de la Dynamique d'Einstein) permettent, connaissant les forces que subit le corpuscule et certaines conditions initiales, de prévoir tout le cours du mouvement. Au corpuscule se trouvent attachées un certain nombre de grandeurs telles que ses coordonnées, son énergie, les composantes de sa quantité de mouvement et celles de son moment de rotation par rapport à un point, etc. ; l'ancienne Mécanique attribue à ces grandeurs une valeur bien déterminée à chaque instant et ses équations permettent de calculer rigoureusement la suite de ces valeurs au cours du temps.

Tout autre est le point de vue de la nouvelle Mécanique. Pour elle, en effet, les grandeurs attachées au corpuscule n'ont pas en général de valeurs bien déterminées de sorte qu'il

n'est plus possible de parler rigoureusement d'une position à chaque instant et d'une trajectoire. On peut seulement assigner à tout instant à toute grandeur attachée à un corpuscule un certain nombre de valeurs possibles, chacune d'elles étant affectée d'une certaine probabilité : ceci veut dire que si l'on effectue à l'instant considéré une mesure précises de la grandeur en question, cette mesure fournira une des valeurs prédites comme possibles, la probabilité pour que l'une de ces valeurs possibles soit le résultat de la mesure pouvant être calculée d'avance.

Ainsi, tandis que le but de la Mécanique ancienne était de prévoir d'une façon rigoureuse et univoque l'évolution à partir d'un état initial donné des grandeurs attachées au corpuscule, le but plus modeste de la Mécanique nouvelle est seulement de calculer les valeurs possibles de ces grandeurs à chaque instant avec leurs probabilités respectives. Au point de vue mathématique, la différence des deux Mécaniques se traduit par le fait suivant : tandis que l'ancienne théorie part d'équations différentielles permettant d'exprimer les coordonnées des points matériels en fonction du temps, la nouvelle part d'une équation aux dérivées partielles ayant la forme d'une équation de propagation d'ondes. Nous allons apprendre à former cette équation fondamentale en nous bornant au cas d'un seul corpuscule placé dans un champ extérieur connu car le cas général d'un système de corpuscules en interaction, facile à traiter dans la Mécanique ondulatoire primitive, ne nous intéresse pas ici, la théorie de Dirac n'étant pas parvenue jusqu'à présent à le transposer d'une façon satisfaisante pour l'électron magnétique.

2. *Formation de l'équation de propagation non relativiste.* — Nous allons écrire l'équation de propagation de la nouvelle Mécanique sous sa forme initiale non relativiste. Cette équation s'obtient en quelque sorte automatiquement en partant de l'expression de l'énergie dans l'ancienne Mécanique newtonienne. Soit un corpuscule de masse m se déplaçant dans

un champ dérivant d'un potentiel U (x, y, z, t). L'expresison classique de l'énergie du corpuscule est :

$$E = \frac{1}{2} m v^2 + U(x, y, z, t) \qquad (1)$$

La quantité de mouvement est par définition le vecteur :

$$\vec{p} = m \vec{v} \qquad (2)$$

de composantes $p_x = m v_x$ etc.

Entre l'énergie et les composantes de la quantité de mouvement, il existe donc la relation :

$$E = \frac{1}{2m} (p_x^2 + p_y^2 + p_z^2) + U(x, y, z, t) \qquad (3)$$

Le second membre de (3) peut se représenter par $H(x, y, z, t, p_x, p_y, p_z)$: c'est la fonction hamiltonienne qui exprime l'énergie à chaque instant t en fonction des coordonnées des corpuscules et des composantes de sa quantité de mouvement (ou moments de Lagrange).

Voici maintenant comment on obtient l'équation de propagation de la Mécanique ondulatoire pour le corpuscule en question. On remplace dans la fonction hamiltonienne p_x par

$$-\frac{h}{2\pi i}\frac{\partial}{\partial x}, \; p_y \text{ par } -\frac{h}{2\pi i}\frac{\partial}{\partial y} \text{ et } p_z \text{ par } -\frac{h}{2\pi i}\frac{\partial}{\partial z};$$

on obtient ainsi un opérateur

$$H\left(x, y, z, t, -\frac{h}{2\pi i}\frac{\partial}{\partial x}, -\frac{h}{2\pi i}\frac{\partial}{\partial y}, -\frac{h}{2\pi i}\frac{\partial}{\partial z}\right)$$

nommé « opérateur Hamiltonien ». On obtient alors l'équation des ondes de la nouvelle Mécanique en écrivant :

$$H(\Psi) = \frac{h}{2\pi i}\frac{\partial \Psi}{\partial t} \qquad (4)$$

$\Psi(x, y, z, t)$ étant la fonction d'onde du corpuscule, fonction essentiellement complexe. En explicitant la forme de l'opérateur H, on trouve aisément l'expression suivante de (4) :

$$\Delta\Psi - \frac{8\pi^2 m}{h^2} U(x, y, z, t)\Psi = \frac{4\pi i m}{h}\frac{\partial \Psi}{\partial t} \qquad (5)$$

Δ étant le symbole Laplacien bien connu. L'équation des ondes étant du premier ordre par rapport au temps permet de calculer la forme de la fonction d'onde à tout instant quand on en connaît la forme à l'instant initial.

Dans le cas très important où U ne dépend pas du temps (champ extérieur constant), l'équation d'ondes admet des solutions monochromatiques, c'est-à-dire ne dépendant du temps que par un facteur de la forme $e^{\frac{2\pi i}{h}Et}$. Une telle onde monochromatique satisfait à l'éqution :

$$\Delta\Psi + \frac{8\pi^2 m}{h^2}[E - U(x, y, z)]\Psi = 0 \qquad (6)$$

ou équation de Schrödinger qui est une forme dégénérée de (5)

Dans le cas plus particulier encore où U est nul (pas de champ extérieur), on peut encore écrire l'équation (6) avec $U = 0$ pour les ondes monochromatiques et l'on a en ce cas comme solution l'onde monochromatique plane :

$$\Psi = a\, e^{\frac{2\pi i}{h}[Et - \sqrt{2mE}(\alpha x + \beta y + \gamma z)]} \qquad (7)$$

a étant l'amplitude constante, α, β, γ étant les cosinus directeurs de la direction de propagation. L'onde (7) a la fréquence $\nu = \frac{E}{h}$ et la longueur d'onde $\lambda = \frac{h}{\sqrt{2mE}} = \frac{h}{mv}$.

C'est là l'onde monochromatique plane que la Mécanique ondulatoire a dès ses débuts fait correspondre au mouvement libre rectiligne et uniforme d'un corpuscule de masse m, d'énergie E et de quantité de mouvement mv.

3. ***Conception nouvelle des grandeurs attachées au corpuscule.*** — Nous venons de voir que le passage de l'ancienne à la nouvelle Mécanique s'opère en remplaçant les composantes de la quantité de mouvement par les opérateurs $-\frac{h}{2\pi i}\frac{\partial}{\partial x}$ etc. Ceci n'est qu'une application particulière d'une idée générale de la nouvelle Mécanique, idée qui consiste à substituer des opérateurs à toutes les grandeurs de la

Dynamique classique. Il est aisé de préciser comment on forme ces opérateurs. Nous savons déjà quels sont les opérateurs correspondant à p_x, p_y, p_z; d'autre part l'opérateur correspondant à l'énergie est l'opérateur Hamiltonien H défini plus haut. A une coordonnée du corpuscule, x par exemple, on fait correspondre l'opérateur x. qui signifie (multiplication par x ». Toutes les autres grandeurs mécaniques sont des grandeurs dérivées de x, y, z, t, p_x, p_y et p_z : donc chaque fois qu'une grandeur s'exprime comme fonction rationnelle entière des coordonnées et des moments, on saura former l'opérateur correspondant ([1]). Ainsi par exemple la composante z du moment de rotation du corpuscule par rapport à l'origine des coordonnées sera remplacée par l'opérateur :

$$M_z = xp_y - yp_x = -\frac{h}{2\pi i}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right) \qquad (8)$$

en prenant des axes directs et en convenant qu'à un mouvement de rotation dans le sens positif sur le plan xy correspond un moment de rotation positif.

Les opérateurs que l'on est amené à former en Mécanique ondulatoire comme correspondant à des grandeurs mécaniques mesurables sont des opérateurs en général complexes appartenant tous à une classe particulière : celle des opérateurs hermitiques.

Voici comment est définie la classe des opérateur hermitiques. Nous allons d'abord faire une convention qui subsistera dans toute la suite de cet ouvrage : une certaine lettre représentant un opérateur ou une fonction, la même lettre affectée d'un astérisque représentera la quantité complexe conjuguée. Cette convention admise, soit A un opérateur du genre de ceux que nous avons considérés; si $d\tau$ désigne l'élément de volume $dx\,dy\,dz$ d'espace, l'opérateur A est par définition hermitique si l'on a :

([1]) Il pourrait y avoir des ambiguités provenant de l'ordre des facteurs, mais nous n'avons pas à nous en occuper ici, car nous n'en rencontrerons pas.

$$\int f^* A(g)\, d\tau = \int g A^* (f^*)\, d\tau \qquad (9)$$

les intégrations étant étendues à tout l'espace et les fonctions f et g des coordonnées étant finies, uniformes et continues dans tout l'espace et tendant vers zéro à l'infini assez vite pour que les intégrales de surface obtenues en intégrant par parties le premier membre de (9) soient nulles. Les opérateurs envisagés en Mécanique ondulatoire sont tous hermitiques ; il est facile de le vérifier pour chaque opérateur et en particulier on le vérifiera par exemple pour l'opérateur M_z de (8).

En dehors de l'hermiticité, les opérateurs de la Mécanique ondulatoire ont aussi toujours un autre caractère commun : ils sont linéaires, c'est-à-dire que l'on a toujours :

$$A(\varphi_1 + \varphi_2) = A(\varphi_1) + A(\varphi_2) \qquad (10)$$

et par suite

$$A(c\,\varphi) = c\,A(\varphi) \qquad (11)$$

Il faut faire entre les opérateurs que la nouvelle Mécanique attache à un corpuscule une distinction importante. Les uns intéressent l'ensemble des 3 coordonnées x, y, z et sont nommés « opérateurs complets ».

Les autres n'intéressent qu'une ou deux coordonnées et sont pour cette raison des « opérateurs incomplets ». Par exemple l'opérateur Hamiltonien H est complet tandis que les opérateurs p_x ou M_z sont incomplets. Nous verrons plus loin l'importance de cette distinction.

Bref, en Mécanique ondulatoire, nous faisons correspondre à toute grandeur dynamique attachée au corpuscule un opérateur linéaire et hermitique. Mais il est bien évident que si l'on fait une mesure précise d'une de ces grandeurs mécaniques, le résultat de la mesure s'exprimera par un nombre réel. Ainsi que nous l'avons dit dans le paragraphe 1, le but de la nouvelle Mécanique est d'abord de nous dire quels sont les nombres réels qu'une mesure précise peut nous fournir comme valeur d'une grandeur mécanique. De l'opérateur

hermitique que la nouvelle Mécanique fait correspondre à une grandeur attachée au corpuscule, nous devons donc pouvoir déduire une liste de nombres réels représentant tous les résultats possibles d'une mesure précise de cette grandeur. Or ceci est possible parce que tous les opérateurs hermitiques de la mécanique ondulatoire possèdent une suite de « valeurs propres » qui sont des nombres réels. C'est ce que nous allons maintenant expliquer.

4. *Valeurs propres et fonctions propres d'un opérateur hermitique.* — Soit A un opérateur linéaire hermitique. Ecrivons l'équation :

$$A(\varphi) = \alpha\varphi \qquad (12)$$

où α est une constante et φ une fonction des coordonnées x, y, z.

Par définition, nous appellerons « valeurs propres de l'opérateur A » les valeurs de la constante α pour lesquelles l'équation (12) admet au moins une solution $\varphi(x, y, z)$ dite « fonction propre de l'opérateur A » jouissant des propriétés suivantes [1] : elle est partout finie, uniforme et continue et l'intégrale du carré de son module dans tout l'espace a un sens et est convergente. Naturellement si l'opérateur A dépend du temps, il en est de même de ses valeurs et fonctions propres.

Nous admettrons l'existence des valeurs propres des opérateurs linéaires et hermitiques de la Mécanique ondulatoire, mais nous allons démontrer que ces valeurs propres sont nécessairement réelles. En effet l'équation conjuguée de (12) s'écrit :

$$A^*(\varphi^*) = \alpha^*\varphi^* \qquad (12^*)$$

et, comme A est linéaire, on a :

$$\int_D \varphi^* A(\varphi)\, d\tau - \int_D \varphi\, A^*(\varphi^*)\, d\tau = (\alpha - \alpha^*) \int_D \varphi\,\varphi^*\, d\tau \qquad (13)$$

---

[1] On ne considère ici comme distinctes que les fonctions linéairement indépendantes.

l'intégration étant étendue à tout le domaine D des variables qui figurent dans φ, c'est-à-dire dans A. Or le premier membre de (13) est nul parce que A est hermitique. Comme l'intégrale du second membre est essentiellement positive, on doit avoir $\alpha = \alpha^*$, donc α est réelle.

L'ensemble des valeurs propres réelles d'un opérateur hermitique est appelé le « spectre » de cet opérateur. Ce spectre est discontinu si les valeurs propres sont isolées et continu si elles forment une suite continue. Un spectre peut même être en partie continu, en partie discontinu. Nous raisonnerons d'abord sur les spectres discontinus.

Désignons par α_i une valeur propre isolée ; il existe au moins une fonction propre $\varphi_i(x, y, z)$ qui lui correspond. L'ensemble des fonctions propres forme un système orthogonal en ce sens que si φ_i et φ_j sont deux fonctions propres correspondant à deux valeurs propres distinctes α_i et α_j, on a :

$$\int_D \varphi_i^* \varphi_j \, d\tau = 0 \tag{14}$$

En effet, α_i étant réelle, on a :

$$A(\varphi_j) = \alpha_j \varphi_j \qquad A^*(\varphi_i^*) = \alpha_i \varphi_i^* \tag{15}$$

et par suite :

$$\int_D \varphi_j A^*(\varphi_i^*) \, d\tau - \int_D \varphi_i^* A(\varphi_j) \, d\tau = (\alpha_i - \alpha_j) \int_D \varphi_i^* \varphi_j \, d\tau \tag{16}$$

Le premier membre étant nul en raison de l'hermiticité de A et $\alpha_i - \alpha_j$ étant par hypothèse différent de zéro, l'équation (14) est démontrée.

La démonstration est en défaut pour deux fonctions propres linéairement indépendantes correspondant à une même valeur propre. Quand ce cas se présente, on dit qu'il y a dégénérescence et que la valeur propre est multiple. Soit α_i une valeur propre multiple à laquelle correspondent p fonctions propres linéairement indépendantes φ_{i1}, φ_{i2} ... φ_{ip}. L'opérateur A étant linéaire, toute combinaison linéaire de φ_{i1} ... φ_{ip} est solution de $A\varphi = \alpha_i\varphi$. On peut donc remplacer les p fonc-

tions propres linéairement indépendantes $\varphi_{i1} \dots \varphi_{ip}$ par p combinaisons linéaires linéairement indépendantes de ces fonctions et l'on voit facilement qu'on peut choisir ces combinaisons linéaires de façon qu'elles soient orthogonales entre elles. En d'autres termes, quand une valeur propre est multiple, le système des fonctions propres linéairement indépendantes n'est déterminé qu'à une transformation linéaire près et l'on peut profiter de cette indétermination partielle pour avoir un système de fonctions propres indépendantes qui soit orthogonal. On peut donc toujours supposer que l'ensemble des fonctions propres d'un opérateur hermitique est orthogonal.

Les fonctions propres d'un opérateur hermitique ne sont déterminées qu'à un facteur constant complexe près (même sans dégénérescence). On a l'habitude, pour fixer le module de ce facteur complexe, de « normer » les fonctions φ_i, c'est-à-dire de poser :

$$\int_D \varphi_i^* \varphi_i \, d\tau = \int_D |\varphi_i|^2 \, d\tau = 1 \tag{17}$$

équation qui a un sens puisque $|\varphi_i|^2$ est sommable. Les fonctions propres une fois normées ont encore un facteur arbitraire de la forme $e^{i\alpha}$.

En introduisant le symoble δ_{ij} égal à un si $i = j$ et à zéro si $i \neq j$, on peut résumer les formules (14) et (17) par la formule :

$$\int_D \varphi_i^* \varphi_j \, d\tau = \delta_{ij} \tag{18}$$

Toutes les formules précédentes s'appliquent au spectre discontinu. Si l'opérateur A possède un spectre continu, à toute valeur propre α de ce spectre correspond une fonction propre que nous écrirons $\varphi(\alpha, x, y, z)$ car, α variant continûment dans le spectre continu, il est plus naturel de l'écrire comme une variable que comme un indice. Les fonctions propres $\varphi(\alpha, x, y, z)$ sont orthogonales aux fonctions propres du spectre discontinu s'il y en a un. Mais, pour éviter cer-

taines difficultés de convergence, il est plus commode dans l'étude des spectres continus de considérer, au lieu des fonctions propres $\varphi(\alpha, x, y, z)$ elles-même, les expressions $\frac{1}{\Delta\alpha}\int_{\alpha}^{\alpha+\Delta\alpha} \varphi(\alpha, x, y, z)\, d\alpha$ dites « différentielles propres » correspondant à des intervalles $\alpha \longrightarrow \alpha + \Delta\alpha$ choisis aussi petits que l'on veut dans le domaine de variation continue de α. L'emploi des différentielles propres conduit à remplacer la formule (18) par la formule :

$$\frac{1}{\Delta\alpha}\int_{D} d\tau \left[\int_{\alpha}^{\alpha+\Delta\alpha} \varphi^{*}(\alpha, x, y, z)\, d\alpha\right]\left[\int_{\alpha'}^{\alpha'+\Delta\alpha} \varphi(\alpha, x, y, z)\, d\alpha\right] = \delta_{\alpha\alpha'}. \tag{19}$$

Avant de clore ce paragraphe, nous devons encore signaler une propriété très importante des fonctions propres d'un opérateur linéaire hermitique A : elles forment un « système complet ». Cela veut dire que, sous des conditions très larges, une fonction des variables intéressées par A (variables du domaine D) se laisse toujours développer en une série de fonctions propres de cet opérateur. Si par exemple $f(x, y, z)$ est une fonction des trois variables x, y, z, elle se laisse très généralement développer suivant les fonctions propres d'un opérateur hermitique complet A sous la forme :

$$f(x, y, z) = \sum_{i} d_i\, \varphi_i(x, y, z) + \int d(\alpha)\, \varphi(\alpha, x, y, z)\, d\alpha \tag{20}$$

la somme Σ étant étendue au spectre discontinu, l'intégrale au spectre continu.

Mettant en évidence les différentielles propres correspondant aux divers intervalles $\Delta\alpha$ du spectre continu, nous pouvons remplacer (20) par :

$$f(x,y,z) = \sum_{i} d_i\, \varphi_i(x,y,z) + \sum_{\Delta\alpha} d(\alpha)\left[\frac{1}{\Delta\alpha}\int_{\alpha}^{\alpha+\Delta\alpha} \varphi(x,y,z)\, d\alpha\right]\Delta\alpha \tag{21}$$

En utilisant les formules (18) et (19), on trouve aisément :

$$d_i = \int \varphi_i^* f(x, y, z)\, d\tau\,;$$

$$d(\alpha) = \int_D \left[\frac{1}{\Delta\alpha}\int_\alpha^{\alpha+\Delta\alpha} \varphi^* (\alpha, x, y, z)\, d\alpha\right] f(x, y, z)\, d\tau\,. \qquad (22)$$

Les quantités d_i et $d(\alpha)$ sont appelées « les coefficients de Fourier » du développement de la fonction $f(x, y, z)$ suivant les fonctions propres de l'opérateur A. La série et l'intégrale de Fourier rentrent comme cas particuliers simples dans ce type de développements.

5. ***Principes généraux de la Mécanique ondulatoire.*** — Nous avons dit au paragraphe 1 de ce chapitre que le but de la nouvelle Mécanique était de calculer les valeurs possibles des grandeurs attachées au corpuscule et leurs probabilités respectives. Nous avons ensuite appris à associer au corpuscule une fonction d'onde $\Psi(x, y, z, t)$ solution de l'équation (5), fonction d'onde que nous supposerons toujours « normée » par la condition :

$$\int \Psi\Psi^*\, d\tau = 1 \qquad (23)\ (^1)$$

Puis nous avons fait correspondre à chaque grandeur attachée au corpuscule un opérateur linéaire hermitique qui permet de définir un ensemble de nombres réels, ses valeurs propres, et un système complet de fonctions normées et orthogonales, ses fonctions propres.

Nous sommes alors en mesure d'énoncer les deux principes fondamentaux suivants de la nouvelle Mécanique :

Premier principe : Les valeurs possibles à l'instant t d'une grandeur attachée au corpuscule, c'est-à-dire les résultats possibles d'une mesure précise faite à l'instant t de cette gran-

---

(1) Nous démontrerons plus loin que si la condition (23) est satisfaite à un certain instant, elle l'est à tout autre instant.

deur, sont les valeurs propres à l'instant t de l'opérateur linéaire hermitique A correspondant à cette grandeur.

Deuxième principe : Si un corpuscule a pour fonction d'onde une certaine solution $\Psi(x, y, z, t)$ de son équation d'ondes, la probabilité pour qu'une mesure précise de la grandeur correspondant à l'opérateur complet A fournisse à l'instant t une certaine valeur propre est égale au carré du module du coefficient de la fonction propre correspondante dans le développement de la fonction d'onde Ψ suivant les fonctions propres normées et orthogonales de l'opérateur A.

Plus explicitement si la fonction Ψ se développe suivant les fonctions propres de A sous la forme [analogue à (21)]:

$$\Psi = \sum_i c_i \varphi_i + \sum_{\Delta\alpha} c(\alpha) \left[\frac{1}{\Delta\alpha} \int_{\alpha}^{\alpha+\Delta\alpha} \varphi(\alpha)\, d\alpha \right] \Delta\alpha \qquad (24)$$

c'est $|c_i|^2$ qui donne la probabilité de la valeur propre α_i et c'est $|c(\alpha)|^2 \Delta\alpha$ qui donne la probabilité d'une valeur comprise dans l'intervalle $\alpha \longrightarrow \alpha + \Delta\alpha$. Comme la fonction Ψ est normée, la probabilité totale de toutes les hypothèses possibles est bien l'unité, on le vérifie aisément. Naturellement les probabilités fournies par le deuxième principe sont en général fonctions du temps t, c'est-à-dire de l'instant de la mesure.

Si l'opérateur A admet des valeurs propres multiples, l'énoncé du second principe doit être complété. Soit α_i une valeur propre multiple à laquelle correspondent p fonctions propres $\varphi_{i1} \dots \varphi_{ip}$ normées et orthogonales, linéairement indépendantes. La probabilité de trouver par une mesure la valeur α_i pour la grandeur en question est

$$|c_{i1}|^2 + |c_{i2}|^2 + \cdots + |c_{ip}|^2,$$

c'est-à-dire la somme des modules des coefficients de $\varphi_i \dots \varphi_{ip}$ dans le développement de Ψ suivant les fonctions propres de A. Cette probabilité est indépendante de la façon dont on a choisi les p fonctions propres $\varphi_{i1} \dots \varphi_{ip}$, comme cela doit être.

Quand l'opérateur A est incomplet, l'énoncé du second principe doit subir une modification. Alors, en effet, les fonctions propres de A ne contiennent pas toutes les trois variables x, y, z et dans le développement (24) les c_i et les $c(\alpha)$ sont évidemment fonctions des variables non contenues dans les φ_i et $\varphi(\alpha)$. La probabilité d'une valeur propre α_i ne peut pas être alors $|c_i|^2$ quantité qui dépendrait encore de certaines variables. Il est donc nécessaire pour obtenir les probabilités d'*intégrer* les expressions indiquées plus haut dans tout le domaine des variables qui ne figurent pas dans A. Par exemple si A dépend seulement de y et de z, les c_i dépendront de x et la probabilité de la valeur α_i sera non pas $|c_i|^2$ mais bien $\int_{-\infty}^{+\infty} |c_i|^2 \, dx$. On vérifiera que cette modification est bien en accord avec l'idée que la probabilité totale de toutes les hypothèses possibles doit être l'unité.

Nous allons indiquer quelques exemples d'application des principes généraux. Un exemple très simple est l'application à l'Hamiltonien qui est, nous le savons, un opérateur complet. Pour l'opérateur Hamiltonien, l'équation (12) s'écrit :

$$H(\varphi) = E\varphi \tag{25}$$

en écrivant E au lieu de α. On a des valeurs propres E_i et des fonctions propres φ_i. Ces valeurs et fonctions propres dépendent du temps si A en dépend, c'est-à-dire si le système n'est pas conservatif. Une mesure précise de l'énergie ne peut fournir comme résultat qu'une des valeurs E_i relatives à l'instant t de la mesure et la probabilité d'obtenir la valeur $E_k(t)$ est égale au carré du module du coefficient de la fonction φ_k dans le développement de la fonction d'ondes Ψ du corpuscule suivant les fonctions propres de l'énergie à l'instant t. C'est là ce que nous avons nommé dans d'autres exposés « le principe de décomposition spectrale ».

Cherchons maintenant à appliquer nos principes à la grandeur: coordonnée x du corpuscule. L'équation (12) prend la forme

$$x \cdot \varphi = \alpha\varphi \tag{26}$$

Cette équation peut être regardée comme satisfaite pour toute valeur réelle de α par la fonction $\delta(x-\alpha)$ ou fonction de Dirac qui jouit des propriétés suivantes : 1° Elle est fonction paire de l'argument $(x-\alpha)$; 2° L'intégrale $\int f(x) \delta(x-\alpha)\, dx$ est nulle si l'intervalle d'intégration ne contenant pas la valeur $x=\alpha$ et est égale à $f(\alpha)$ pour toute valeur d'intégration contenant cette valeur. L'équation (26) admet donc un spectre continu comprenant toutes les valeurs réelles de α de $-\infty$ à $+\infty$.

D'après le premier principe, une mesure de la coordonnée peut donc donner *a priori* (comme cela doit être) n'importe quelle valeur de $-\infty$ à $+\infty$. De plus les différentielles propres $\frac{1}{\Delta\alpha}\int_{\alpha}^{\alpha+\Delta\alpha} \delta(x-\alpha)\, d\alpha$ de ce spectre continu forment comme on le voit aisément un système complet normé et orthogonal. Comme l'on a d'après la définition des fonctions δ :

$$\Psi(x, y, z, t) = \int_{-\infty}^{+\infty} \Psi(\alpha, y, z, t)\, \delta(x-\alpha)\, d\alpha \tag{27}$$

la probabilité pour qu'une mesure de la coordonnée x donne une valeur comprise dans l'intervalle $\alpha \rightarrow \alpha + d\alpha$ est d'après le second principe :

$$\int_{-\infty}^{+\infty} dz \int_{-\infty}^{+\infty} dy\, |\Psi(\alpha, y, z, t)|^2\, d\alpha \tag{28}$$

Il en résulte facilement que la probabilité pour qu'une mesure simultanée des trois coordonnées donne des valeurs comprises dans les intervalles $\alpha \rightarrow \alpha + d\alpha$, $\beta \rightarrow \beta + d\beta$, $\gamma \rightarrow \gamma + d\gamma$ est $|\Psi(\alpha, \beta, \gamma, t)|^2\, d\alpha\, d\beta\, d\gamma$; ou, ce qui revient au même, la probabilité pour qu'une mesure permette de localiser le corpuscule dans élément de volume $dx\, dy\, dz$ autour du point de coordonnées x, y, z est $|\Psi(x, y, z, t)|^2\, dx\, dy\, dz$. C'est là ce que nous avons appelé dans d'autres exposés le principe des interférences.

6. ***Grandeurs simultanément ou non simultanément mesurables.*** — Des principes généraux énoncés aux paragraphes précédents, on déduit une conséquence très importante : deux grandeurs mécaniques ne peuvent être simultanément mesurées avec précision que si les opérateurs correspondants A et B permutent, c'est-à-dire si l'on a $AB = BA$.

En effet, si φ_i et χ_i désignant les fonctions propres de A et de B respectivement, α_i et β_i leurs valeurs propres, pour qu'une mesure simultanée des deux grandeurs en question puisse être faite avec précision, il faut que l'on puisse attribuer simultanément avec certitude à la première grandeur une certaine valeur α_i et à la seconde une certaine valeur β_i ; d'après le second principe il faut donc que l'on puisse écrire la fonction d'onde Ψ du corpuscule sous la forme :

$$\Psi = c_i \varphi_i = d_i \chi_i \tag{29}$$

c_i peut dépendre des variables qui ne figurent pas dans φ_i si A est un opérateur incomplet et d_i peut de même dépendre des variables qui ne figurent pas dans χ_i si β est incomplet. De l'équation précédente, on tire :

$$AB(\Psi) = AB(d_i \chi_i) = A(d_i \beta_i \chi_i) = \beta_i A(c_i \varphi_i) = \beta_i \alpha_i \Psi \tag{30}$$

et

$$BA(\Psi) = BA(c_i \varphi_i) = B(c_i \alpha_i \varphi_i) = \alpha_i B(d_i \chi_i) = \alpha_i \beta_i \Psi. \tag{31}$$

On doit donc avoir

$$AB(\Psi) = BA(\Psi) \tag{32}$$

pour toute valeur de Ψ de la forme (29), ce qui entraîne $AB = BA$.

L'exemple le plus simple et le plus important de grandeurs qui ne sont pas simultanément mesurables est celui d'une coordonnée et de la composante correspondante de la quantité de mouvement. On a en effet :

$$xp_x - p_x x = \frac{h}{2\pi i}\left(\frac{\partial}{\partial x} x \cdot - x \cdot \frac{\partial}{\partial x}\right) = \frac{h}{2\pi i}. \tag{33}$$

Autrement dit l'opérateur $xp_x - p_x x$ est équivalent à la multiplication par $\frac{h}{2\pi i}$. Donc x. et p_x ne permutent pas et par suite une coordonnée et la composante correspondante de quantité de mouvement ne peuvent être simultanément mesurées avec précision. Une coordonnée et le moment de Lagrange correspondant ne sont donc jamais connus à un instant donné qu'avec certaines incertitudes Δx et Δp_x qui ne peuvent être nulles à la fois toutes les deux. On peut démontrer que l'on a toujours :

$$\Delta x \,.\, \Delta p_x \geqslant h \tag{34}$$

du moins en ordre de grandeur. L'inégalité (34) et les deux inégalités analogues pour y et z constituent les « relations d'incertitude » d'Heisenberg sur lesquelles nous avons longuement insisté dans d'autres ouvrages (1).

(1) *Introduction à l'étude de la Mécanique ondulatoire*, Hermann, Paris 1930.

La théorie de la quantification dans la nouvelle Mécanique, Hermann, Paris, 1932.

On trouvera dans ces deux ouvrages un exposé plus détaillé des principes exposés sommairement dans le présent chapitre et dans le chapitre suivant.

CHAPITRE VI

Résumé des principes de la mécanique ondulatoire (*suite*)

1. *Quelques définitions relatives aux matrices algébriques*

ON appelle « matrice » un tableau de nombres comportant un nombre fini ou infini de lignes et de colonnes. Si le tableau est de dimensions finies, nous le supposerons de forme carrée ; on pourrait le supposer rectangulaire pour plus de généralité, mais ce serait pour nous ici une complication inutile. Chaque nombre figurant dans le tableau (élément de la matrice) peut être repéré à l'aide de deux indices marquant respectivement la ligne et la colonne auxquelles il appartient. Désignons donc par a_{ik} l'élément de la matrice qui se trouve inscrit dans le tableau à l'intersection de la i^e ligne et de la k^e colonne : l'ensemble de la matrice sera alors représenté par A ou $|a_{ik}|$. Les éléments a_{ii} à indices égaux sont situés sur la diagonale du tableau et sont appelés « éléments diagonaux ». Nous dirons que deux matrices A et B sont égales et nous écrirons $A = B$ si leurs éléments de mêmes indices sont tous égaux ($a_{ik} = b_{ik}$).

Les matrices se présentent en algèbre quand on étudie les transformations linéaires. En effet si des variables x_i' sont des combinaisons linéaires d'autres variables x_i, on a des formules de transformation du type :

$$x_i' = \sum_j a_{ij}\, x_j \tag{1}$$

formules qu'on peut condenser en écrivant la relation vectorielle :

$$X' = AX \tag{2}$$

et en convenant que le vecteur AX a pour composante d'indice i la quantité $(AX)_i = \sum_j a_{ij}\, x_j$.

La formule (1) conduit à définir la somme et le produit de deux matrices ayant même nombre de lignes et de colonnes par les conventions suivantes :

1° La somme de deux matrices A et B est la matrice $A + B$ dont l'élément d'indices ik est $a_{ik} + b_{ik}$;

2° Le produit de la matrice B par la matrice A est la matrice AB dont l'élément ik est $(AB)_{ik} = \sum_l a_{il} \cdot b_{lk}$.

De la définition 2° résulte qu'en général la matrice produit BA n'est pas égale à la matrice produit AB. On dit qu'en général deux matrices ne permutent pas ou ne commutent pas. Si $AB = -BA$, les matrices anti-commutent.

Les éléments d'une matrice peuvent être réels ou complexes. Plaçons-nous dans le cas général des matrices à éléments complexes. Les formules de transformation (1) expriment alors qu'on passe de certaines variables complexes x_i à d'autres variables complexes x_i'. Nous allons maintenant définir quelques types particuliers très importants de matrices complexes.

Nous dirons qu'une matrice est *hermitique* si les éléments symétriques par rapport à la diagonale sont complexes conjugués $(a_{ik} = a_{ki}^*)$. Les éléments diagonaux d'une matrice hermitique sont réels. Si tous les éléments d'une matrice hermitique sont réels, la matrice est symétrique par rapport à sa diagonale. Nous dirons qu'une matrice est anti-hermitique si l'on a $a_{ik} = - a_{ki}^*$. Les éléments diagonaux d'une matrice anti-hermitique sont purement imaginaires. Le produit de deux matrices hermitiques n'est hermitique que si elles com-

mutent ; si elles anti-commutent, le produit est anti-hermitique.

On appelle « matrice adjointe de la matrice A » et l'on désigne par A^+ la matrice obtenue à partir de A en permutant les termes symétriques par rapport à la diagonale et en prenant les quantités complexes conjuguées ; on a donc $a_{ik}^+ = a_{ki}^*$. De cette définition résulte qu'une matrice hermitique est égale à son adjointe : si A est hermitique, on a $A = A^+$. On démontre aisément la formule $(AB)^+ = B^+ A^+$ et il est évident que $(A^+)^+ = A$.

Une matrice est dite « diagonale » quand seuls ses éléments diagonaux sont différents de zéro. Une matrice hermitique et diagonale très importante est la matrice unité que l'on représente par 1 : c'est la matrice dont l'élément *ik* est égal à δ_{ik}.

Etant donnée une matrice A, s'il existe une autre matrice A^{-1} telle que $A . A^{-1} = A^{-1} A = 1$, la matrice A^{-1} est dite la « matrice inverse de A ». Cette matrice inverse, si elle existe, est toujours unique. Quand A a un nombre fini de lignes et de colonnes, la matrice inverse existe toujours si le déterminant formé à l'aide du tableau des a_{ik} n'est pas nul. Quand A a un nombre infini de lignes et de colonnes, A^{-1} peut ne pas exister et il faut dans chaque cas particulier vérifier son existence. On vérifiera aisément la formule : $(AB)^{-1} = B^{-1} A^{-1}$.

Quand A est une matrice à éléments réels et que l'on a :

$$\sum_i a_{ij}\, a_{ik} = \delta_{jk} \qquad (3)$$

on dit que la matrice A est orthogonale : elle définit une transformation orthogonale qui laisse invariante la quantité $\sum_i x_i^2$. Ceci est bien connu. Pour les matrices à éléments complexes, on peut généraliser cette définition : si A est une matrice complexe et si l'on a :

$$\sum_i a_{ij}\, a_{ik}^* = \delta_{jk} \qquad (4)$$

on dit que A définit une transformation complexe orthogo-

nale ou encore qu'elle est « unitaire ». Pour une transformation complexe orthogonale, la quantité $\sum_i x_i^* x_i$ reste invariante. La condition (4) peut d'ailleurs s'écrire

$$\sum_i a_{ki}^+ a_{ij} = \delta_{kj} \quad \text{ou} \quad A^+A = 1 \; ; \; A^+ = A^{-1} \tag{5}$$

Donc pour qu'une matrice soit unitaire, son adjointe doit coïncider avec son inverse.

Soit encore A une matrice et S une matrice unitaire ayant même nombre de lignes et colonnes. La matrice $B = S^{-1} AS$ est dite obtenue à partir de A par une « transformation canonique ». Si A est hermitique, B est aussi hermitique. En effet comme par hypothèse $S^+ = S^{-1}$ et $A^+ = A$, on a :

$$B^+ = (S^{-1} AS)^+ = S^+ A^+ (S^{-1})^+ = S^{-1} AS = B \,. \tag{6}$$

D'où le théorème important : une transformation canonique transforme une matrice hermitique en une autre matrice hermitique [1].

2 *Matrices de la Mécanique ondulatoire.* — Supposons que nous connaissions un système complet de fonctions normées et orthogonales : $\varphi_1 \ldots \varphi_n \ldots$; nous les appellerons les fonctions de base. Un tel système nous est fourni par exemple par l'ensemble des fonctions propres d'un opérateur hermitique.

Etant donné ce système de base, à tout opérateur linéaire nous pouvons faire correspondre une matrice. Soit en effet A un opérateur linéaire ; l'application de cet opérateur à une des fonctions de base φ_i nous fournit une fonction qui doit pouvoir se développer suivant le système complet $\varphi_1 \ldots \varphi_n \ldots$ Nous avons donc une relation de la forme :

$$A(\varphi_i) = \sum_j a_{ji} \varphi_j \tag{7}$$

d'où en raison des propriétés des φ_i

[1] Cet énoncé suppose essentiellement que la matrice S de la transformation canonique est unitaire.

$$a_{ji} = \int_D \varphi_j^* \, A(\varphi_i) \, d\tau \tag{8}$$

D étant le domaine des variables qui figurent dans les φ_i.

Par définition les a_{ji} de la formule (8) sont les éléments de la matrice engendrée par l'opérateur A dans le système de base des φ_i. Nous désignerons aussi cette matrice par la lettre A ; si nous tenons à préciser le système des fonctions de base employé, nous pourrons la désigner par A^φ.

Les matrices ainsi obtenues peuvent être nommées « matrices de la Mécanique ondulatoire ». Nous allons vérifier qu'elles satisfont bien aux règles d'addition et de multiplication des matrices algébriques. Pour cela considérons deux opérateurs linéaires A et B. Nous aurons :

$$A(\varphi_i) = \sum_j a_{ji} \varphi_j \qquad B(\varphi_i) = \sum_j b_{ji} \varphi_j \tag{9}$$

d'où

$$(A + B)(\varphi_i) = \sum_j (a_{ji} + b_{ji}) \varphi_j \tag{10}$$

L'élément *ji* de la matrice A + B est donc $a_{ji} + b_{ji}$; c'est bien la règle d'addition des matrices algébriques. On a de plus :

$$AB(\varphi_i) = \sum_j b_{ji} \, A(\varphi_j) = \sum_j b_{ji} \sum_k a_{kj} \varphi_k = \sum_k \left(\sum_j a_{kj} b_{ji} \right) \varphi_k \tag{11}$$

L'élément *ki* de la matrice AB est donc $\sum_j a_{kj} b_{ji}$; c'est bien la règle de multiplication des matrices algébriques.

La condition pour qu'une matrice A de la Mécanique ondulatoire soit hermitique est :

$$a_{ji} = \int_D \varphi_j^* \, A(\varphi_i) \, d\tau = a_{ij}^* = \int_D \varphi_i \, A^*(\varphi_j^*) \, d\tau \, . \tag{12}$$

Or nous savons que si cette condition est remplie pour toutes les fonctions φ_i, l'opérateur A est par définition hermitique et inversement. Donc la condition nécessaire et suffisante pour qu'une matrice de la Mécanique ondulatoire soit hermitique est que l'opérateur dont elle dérive le soit. L'her-

miticité est donc une propriété intrinsèque de l'opérateur en ce sens qu'un opérateur hermitique engendre dans tout système de fonctions de base une matrice hermitique.

Tous les opérateurs que nous considérons en Mécanique ondulatoire étant hermitiques, les matrices qui leur correspondent le sont également.

3. ***Valeurs moyennes en Mécanique ondulatoire.*** — Envisageons un corpuscule et supposons connue la fonction d'onde Ψ qui lui est associée. Soit d'autre part une des grandeurs mécaniques attachées au corpuscule à laquelle correspond dans la nouvelle Mécanique l'opérateur A ; en parlant de cette grandeur nous dirons pour abréger « la grandeur A ».

Les principes généraux énoncés dans le dernier chapitre nous permettent de prévoir les grandeurs possibles de la grandeur A et leurs probabilités respectives. Comme il y a en général plusieurs valeurs possibles de probabilités non nulles, on ne peut parler sans équivoque de la valeur de la grandeur A à chaque instant, mais on peut parler de sa valeur moyenne, cette valeur moyenne étant définie de la façon usuelle comme la somme des produits de chaque valeur possible par la probabilité correspondante. Si α_i et φ_i désignent les valeurs propres et fonctions propres de l'opérateur A et si la fonction d'onde Ψ admet le développement :

$$\Psi = \sum_i c_i \varphi_i \tag{13}$$

la valeur moyenne $\overline{A}$ est d'après les principes généraux :

$$\overline{A} = \sum_i \alpha_i |c_i|^2 \tag{14}$$

Ceci peut s'écrire sous la forme essentielle en Mécanique ondulatoire :

$$\overline{A} = \int_D \Psi^* A(\Psi)\, d\tau . \tag{15}$$

L'équivalence de (14) et de (15) résulte de la formule :

$$\int_D \Psi^* A(\Psi)\, d\tau = \int_D \sum_i c_i^* \varphi_i^* \cdot A\left(\sum_k c_k \varphi_k\right) d\tau$$

$$= \sum_{ik} c_i^* c_k \alpha_k \int_D \varphi_i^* \varphi_k \, d\tau \qquad (16)$$

et du fait que les fonctions φ_i sont normées et orthogonales. La valeur moyenne $\overline{A}$ définie par (14) et (15) est évidemment une quantité réelle.

Le raisonnement par lequel nous venons d'établir la formule fondamentale (15) n'est rigoureusement valable que pour les opérateurs A complets sans spectre continu et sans valeurs propres multiples. Mais il est facile de l'étendre aux opérateurs incomplets ou dégénérés et aux spectres continus : la formule (15) est générale.

La forme de l'expression (15) de $\overline{A}$ nous permet de dire que $\Psi^* A(\Psi)$ est la « densité de valeur moyenne » pour la grandeur A. Mais cette « densité » est d'une nature bien différente de celles qu'on a à considérer dans les théories classiques ; en effet l'élément d'intégration $\Psi^* A(\Psi) d\tau$ (qui est même généralement complexe) ne peut aucunement être considéré comme une certaine quantité de la grandeur A qui serait localisée dans l'élément $d\tau$ et seule l'intégrale (15), qui est toujours réelle, a un sens physique. Ceci est une remarque importante qu'il faut garder présente à l'esprit.

La formule (15) fournit une interprétation statistique des matrices de la Mécanique ondulatoire ou du moins de leurs éléments diagonaux. Nous allons le montrer en raisonnant encore sur des opérateurs complets non dégénérés et sans spectre continu car le raisonnement général ne comporte que quelques complications sans modification du résultat. Supposons que le développement de la fonction d'ondes Ψ suivant les fonctions propres φ_i de l'opérateur A se réduise à un seul terme ; on a alors

$$\Psi = c_i \varphi_i \qquad (17)$$

avec $|c_i| = 1$ puisque Ψ est toujours supposée normée. Nous sommes en ce cas sûrs qu'une mesure de la grandeur A fournira la valeur α_i. Soit maintenant une autre grandeur B atta-

chée au corpuscule et correspondant à un opérateur B. Nous obtenons la valeur moyenne de la grandeur B par application de la formule (15) et nous trouvons :

$$\overline{B} = \int_D \Psi^* B(\Psi)\, d\tau = \int_D \varphi_i^* B(\varphi_i)\, d\tau . \qquad (18)$$

Or la deuxième intégrale de (18) n'est pas autre chose que l'élément diagonal d'indice $i\,i$ de la matrice engendrée par l'opérateur A dans le système des fonctions φ_i. D'où le théorème : « L'élément diagonal d'indice $i\,i$ de la matrice engendrée par l'opérateur B dans le système des fonctions propres de l'opérateur A est égale à la valeur moyenne de la grandeur B *quand on sait que la grandeur* A *a la valeur* α_i. »

4. *Valeur moyenne d'une coordonnée. Théorème d'Ehrenfest.* — Considérons une des coordonnées d'un corpuscule, la coordonnée x par exemple. Sa valeur moyenne d'après la formule (15) et en accord avec le principe des interférences est :

$$\overline{x} = \int_D x \Psi^* \Psi \, d\tau . \qquad (19)$$

C'est donc la coordonnée x d'un fluide *fictif* dont la densité serait donnée en chaque point par :

$$\rho = \Psi \Psi^* \qquad (20)$$

Nous appellerons ce fluide fictif le « fluide de probabilité »; la quantité de ce fluide contenue dans un élément de volume $d\tau$ est $\Psi\Psi^* d\tau$ [1] et la quantité totale du fluide reste constante au cours du temps, égale à 1 si la fonction Ψ est normée. Comme la répartition du fluide de probabilité varie au cours du temps, nous lui attribuerons en chaque point au cours du temps une vitesse que nous définirons par la formule :

$$\overrightarrow{u} = \frac{1}{\Psi\Psi^*} \frac{h}{4\pi i m} \left[\Psi \,\mathrm{grad}\, \Psi^* - \Psi^* \,\mathrm{grad}\, \Psi \right] \qquad (21)$$

---

(1) D'après le principe des interférences, cette quantité est donc égale à la « probabilité de présence » du corpuscule dans l'élément $d\tau$.

Il est facile de montrer que dans ces conditions le fluide de probabilité se conserve au cours du temps et cette démonstration établit accessoirement la proposition suivante dont nous avions admis l'exactitude : « Si la fonction Ψ est normée à un instant donné, elle reste normée à toute époque. »

Voici la démonstration en question. De l'équation d'ondes à laquelle obéit la fonction Ψ (équation (5) du chapitre précédent) et de l'équation conjuguée, on tire aisément :

$$\frac{\partial}{\partial t}(\Psi\Psi^*) = \frac{h}{4\pi im}(\Psi^*\Delta\Psi - \Psi\Delta\Psi^*)$$

$$= -\frac{h}{4\pi i m}\sum_{x,y,z}\left[\frac{\partial}{\partial x}\left(\Psi\frac{\partial\Psi^*}{\partial x} - \Psi^*\frac{\partial\Psi}{\partial x}\right)\right] \qquad (22)$$

ou encore d'après les définitions (20) et (21) :

$$\frac{\partial\rho}{\partial t} + \text{div}\,(\rho\,\overrightarrow{u}) = 0\,. \qquad (23)$$

Or l'équation (23) est l'équation hydrodynamique de continuité qui exprime précisément la conservation du fluide dont la répartition et le mouvement sont définis par (20) et (21).

Faisons maintenant une remarque. D'après notre définition des valeurs moyennes, si $f(x, y, z, t)$ est une certaine fonction scalaire ou vectorielle, nous devons appeler « valeur moyenne de cette fonction à l'instant t » la quantité :

$$\bar{f}(t) = \int_D \Psi^* \cdot f(x, y, z, t) \cdot \Psi\, d\tau\,. \qquad (24)$$

Cela étant nous pouvons énoncer un théorème important dû à M. Ehrenfest :

Théorème : Le point de coordonnées $\bar{x}$, $\bar{y}$, $\bar{z}$ se déplace au cours du temps comme le ferait d'après les lois de la Mécanique classique un point matériel de masse m sous l'action d'une force égale à chaque instant à la valeur moyenne $\bar{f}(t)$ de la force réelle.

On a d'abord d'après (19) en employant (22) et en intégrant par parties :

$$\frac{d\bar{x}}{dt}=\int_D x\,\frac{\partial(\Psi\Psi^*)}{\partial t}d\tau=-\frac{h}{4\pi im}\int_D x\sum_{x,y,z}\frac{\partial}{\partial x}\left[\Psi\frac{\partial\Psi^*}{\partial x}-\Psi^*\frac{\partial\Psi}{\partial x}\right]d\tau$$

$$=\frac{h}{4\pi im}\int_D\left(\Psi\frac{\partial\Psi^*}{\partial x}-\Psi^*\frac{\partial\Psi}{\partial x}\right)d\tau=\int_D u_x\,\Psi^*\,\Psi\,d\tau=\bar{u}_x\,. \qquad (25)$$

On trouve ensuite :

$$\frac{d^2\bar{x}}{dt}=\frac{h}{4\pi im}\int_D\left[\frac{\partial\Psi}{\partial t}\frac{\partial\Psi^*}{\partial x}-\frac{\partial\Psi}{\partial x}\frac{\partial\Psi^*}{\partial t}+\Psi\frac{\partial^2\Psi^*}{\partial x\,\partial t}-\Psi^*\frac{\partial^2\Psi}{\partial x\partial t}\right]d\tau$$

$$=\frac{h}{2\pi im}\int_D\left(\frac{\partial\Psi}{\partial t}\frac{\partial\Psi^*}{\partial x}-\frac{\partial\Psi}{\partial x}\frac{\partial\Psi^*}{\partial t}\right)d\tau \qquad (26)$$

ce qui donne en vertu de l'équation de propagation :

$$\frac{d^2\bar{x}}{dt^2}=-\frac{h^2}{8\pi^2m^2}\int_D\left[\frac{\partial\Psi^*}{dx}\left(\Delta\Psi-\frac{8\pi^2m}{h^2}\mathrm{U}\Psi\right)\right.$$

$$\left.+\frac{\partial\Psi}{\partial x}\left(\Delta\Psi^*-\frac{8\pi^2m}{h^2}\mathrm{U}\Psi^*\right)\right]d\tau\,. \qquad (27)$$

Or le théorème de Green donne (après intégration par parties) :

$$\int_D\left(\frac{\partial\Psi^*}{\partial x}\Delta\Psi+\frac{\partial\Psi}{\partial x}\Delta\Psi\right)d\tau=\int_D\left[\frac{\partial\Psi}{\partial x}\Delta\Psi-\Psi^*\frac{\partial}{\partial x}(\Delta\Psi)\right]dt=0 \qquad (28)$$

de sorte que (27) nous donne :

$$m\,\frac{d^2\bar{x}}{dt^2}=\int_D \mathrm{U}\,\frac{\partial}{\partial x}(\Psi\Psi^*)\,d\tau$$

$$=\int_D\left(-\frac{\partial\mathrm{U}}{\partial x}\right)\Psi\Psi^*\,d\tau=-\overline{\frac{\partial\mathrm{U}}{\partial x}}=\bar{f_x}\,. \qquad (29)$$

C'est bien l'expression du théorème d'Ehrenfest pour $\bar{x}$ et l'on établit de même les formules correspondantes pour $\bar{y}$ et $\bar{z}$.

Pour terminer ce paragraphe, nous donnerons encore la définition des densités moyennes de charge électrique et de

courant électrique correspondant à un corpuscule de charge ε dont on connaît la fonction Ψ. La charge ε étant une grandeur physique susceptible d'une seule valeur est égale à sa valeur moyenne et l'on peut écrire :

$$\bar{\varepsilon} = \varepsilon = \int_D \varepsilon \Psi\Psi^* \, d\tau . \tag{30}$$

On peut donc considérer la quantité

$$\delta = \varepsilon \Psi\Psi^* \tag{31}$$

comme la densité moyenne de charge électrique associée au corpuscule. D'autre part, au point de vue classique, le corpuscule de charge ε en mouvement avec une vitesse $\vec{v}$ est équivalent à un élément de courant $\vec{i} = \varepsilon \vec{v}$. Ici nous aurons à remplacer v_x par $\frac{1}{m} p_x$, etc. et les trois composantes $i_x\, i_y\, i_z$ du courant en question correspondront aux opérateurs

$$-\frac{h}{2\pi i}\frac{\varepsilon}{m}\frac{\partial}{\partial x}, \quad -\frac{h}{2\pi i}\frac{\varepsilon}{m}\frac{\partial}{\partial y}, \quad -\frac{h}{2\pi i}\frac{\varepsilon}{m}\frac{\partial}{\partial z}.$$

La valeur moyenne de i_x par exemple sera, toujours d'après (15) :

$$\bar{i}_x = -\frac{h\varepsilon}{2\pi i m}\int_D \Psi^* \frac{\partial \Psi}{\partial x} d\tau = \frac{h\varepsilon}{4\pi i m}\int_D \left(\Psi \frac{\partial \Psi^*}{\partial x} - \Psi^* \frac{\partial \Psi}{\partial x}\right) d\tau \tag{32}$$

et l'on aura des expressions analogue pour $\bar{i}_y$ et $\bar{i}_z$. De ces expressions résulte que la quantité vectorielle :

$$\vec{j} = \frac{h\varepsilon}{4\pi i m}\left[\Psi \operatorname{grad} \Psi^* - \Psi^* \operatorname{grad} \Psi\right] \tag{33}$$

peut être considérée comme la densité moyenne de courant électrique associée au corpuscule. En comparant les formules (30) et (33) aux formules (20) et (21), on voit que les densités δ et $\vec{j}$ sont les densités de charge et de courant qui existeraient d'après la théorie classique si la charge ε du corpuscule était répartie dans le fluide de probabilité proportionnellement à la densité $\Psi\Psi^*$ de celui-ci.

5. ***Intégrales premières en Mécanique ondulatoire.*** — En Mécanique classique, on appelle « intégrale première » une grandeur mécanique s'exprimant à l'aide des coordonnées, des moments et parfois du temps qui reste constante au cours du mouvement en vertu même des équations du mouvement.

Comment définira-t-on une intégrale première dans la nouvelle Mécanique? Voici la réponse qu'on doit donner à cette question : si une grandeur mécanique correspond à un opérateur A, cette grandeur est intégrale première dans un problème déterminé (c'est-à-dire pour une forme donnée de l'Hamiltonien H) si l'on a :

$$\frac{\partial A}{\partial t} + \frac{2\pi i}{h}(AH - HA) = 0 \qquad (34)$$

où $\frac{\partial A}{\partial t}$ représente l'opérateur que l'on obtient en dérivant formellement l'expression de A par rapport à la variable t. Si A ne dépend pas de t, $\frac{\partial A}{\partial t}$ est nul et la condition (34) exprime simplement que A commute avec H. On démontre d'ailleurs que, si H est indépendant du temps, la condition (34) a la signification suivante : les éléments de la matrice engendrée par l'opérateur A dans le système des fonctions propres de l'opérateur H sont constantes.

Si le champ est constant, H ne dépendant pas du temps, l'énergie est visiblement intégrale première : on retrouve un théorème classique. Si la composante x du champ est nulle, H ne dépend pas de x et permute avec p_x : la composante x de la quantité de mouvement est intégrale première comme en Mécanique classique, etc.

Le cas le plus intéressant pour nous sera celui du moment de rotation. Quand le champ présente la symétrie cylindrique autour d'un axe oz, H ne dépend pas de l'azimut φ autour de cet axe. En prenant des axes directs et en convenant qu'à une rotation de sens positif dans le plan xy correspond un moment de rotation positif, l'opérateur correspondant au moment de rotation autour de oz est :

$$M_z = xp_y - yp_x = \frac{h}{2\pi i}\left(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\right) \tag{35}$$

En prenant des coordonnées sphériques dont *oz* soit l'axe polaire, on trouve :

$$M_z = -\frac{h}{2\pi i}\frac{\partial}{\partial \varphi} \tag{36}$$

Par suite, M_z indépendant de t permute avec H et est intégrale première. Si le champ avait la symétrie sphérique autour de 0, chacun des moments de rotation $M_x M_y M_z$ serait intégrale première.

Nous pouvons simplifier la forme de la condition (34) en introduisant l'opérateur :

$$L = H - \frac{h}{2\pi i}\frac{\partial}{\partial t} \tag{37}$$

Comme on a, f étant une fonction quelconque :

$$\frac{\partial}{\partial t}\cdot A(f) = \frac{\partial A}{\partial t}(f) + A\left(\frac{\partial f}{\partial t}\right)$$

l'opérateur $\frac{\partial A}{\partial t}$ est équivalent à l'opérateur $\frac{\partial}{\partial t}\cdot A - A\cdot\frac{\partial}{\partial t}$.

La condition (34) peut donc s'écrire :

$$\begin{aligned} 0 &= (AH - HA) + \frac{h}{2\pi i}\left(\frac{\partial}{\partial t}\cdot A - A\cdot\frac{\partial}{\partial t}\right) \\ &= A\left(H - \frac{h}{2\pi i}\frac{\partial}{\partial t}\right) - \left(H - \frac{h}{2\pi i}\frac{\partial}{\partial t}\right)A. \end{aligned} \tag{38}$$

ou simplement :

$$AL - LA = 0 \tag{39}$$

La condition pour que la grandeur A soit intégrale première est donc simplement que l'opérateur A permute avec L.

CHAPITRE VII

Forme relativiste de la Mécanique ondulatoire à une fonction d'onde

1. *Rappel de quelques formules de la Mécanique relativiste*

L'INTRODUCTION du principe de Relativité en Mécanique a conduit Einstein, bien avant l'apparition de la nouvelle Mécanique, à modifier les équations classiques de la Dynamique Newtonienne. Cette modification n'entraînait d'ailleurs qu'un simple changement de forme de certaines formules ; la Mécanique relativiste d'Einstein conservait toutes les conceptions classiques de point matériel, de vitesse, de trajectoire, de déterminisme mécanique, etc. Comparée à la Mécanique nouvelle, la Dynamique d'Einstein apparaît donc seulement comme une légère modification de la théorie classique ayant pour objet de la rendre conforme au principe de Relativité. Dans le paragraphe 2 du chapitre V, nous sommes partis de formules de la Mécanique Newtonienne pour obtenir l'équation de propagation de la Mécanique ondulatoire ; nous avons ainsi obtenu une Mécanique ondulatoire qui naturellement n'est pas relativiste. Pour obtenir une Mécanique ondulatoire relativiste, il paraît bien naturel d'opérer comme au paragraphe 2 du chapitre mais en partant des formules de la théorie Einsteinienne ; pour cela

nous commencerons par rappeler quelques-unes de ces formules.

En Mécanique relativiste, chaque corpuscule est caractérisé par une grandeur invariante m_0, sa masse propre. Un des principes fondamentaux de la théorie de Relativité étant la proportionnalité de la masse et de l'énergie, le corpuscule de masse m_0 possédera même au repos une énergie interne ou « énergie propre » donnée par :

$$W_0 = m_0 c^2 \tag{1}$$

où c est la vitesse de la lumière dans le vide. Si le corpuscule est en mouvement avec une vitesse $v = \beta c$, son énergie est :

$$W = \frac{m_0 c^2}{\sqrt{1-\beta^2}}. \tag{2}$$

Il y a encore proportionnalité entre l'énergie et la masse à condition de considérer que par suite du mouvement la masse augmente et devient $\frac{m_0}{\sqrt{1-\beta^2}}$.

On peut appeler « énergie cinétique » du corpuscule animé de la vitesse βc la quantité :

$$T = \frac{m_0 c^2}{\sqrt{1-\beta^2}} - m_0 c^2 = m_0 c^2 \left(\frac{1}{\sqrt{1-\beta^2}} - 1 \right) \tag{3}$$

qui représente l'augmentation d'énergie due au mouvement.

Les formules précédentes sont valables en l'absence de champ. Si le corpuscule est soumis à un champ de force dérivant de la fonction potentielle $U(x, y, z, t)$, on doit écrire au lieu de (2) :

$$W = \frac{m_0 c^2}{\sqrt{1-\beta^2}} + U(x, y, z, t) \tag{4}$$

Si β est petit devant l'unité ($v \ll c$), on obtient en première approximation :

$$W = m_0 c^2 + \frac{1}{2} m v^2 + U(x, y, z, t) \tag{5}$$

et en posant

$$E = W - m_0 c^2 \tag{6}$$

on retombe sur la formule classique ; E est donc l'énergie au sens de la Mécanique classique qui diffère de l'énergie W de la Mécanique relativiste par le terme constant d'énergie propre $m_0 c^2$.

La quantité de mouvement du corpuscule de masse m_0 animé de la vitesse $v = \beta c$ est dans la théorie d'Einstein :

$$\vec{p} = \frac{m_0}{\sqrt{1-\beta^2}} \vec{v}\,. \tag{7}$$

Elle est en somme égale au produit de la vitesse par la masse en mouvement $\frac{m_0}{\sqrt{1-\beta^2}}$.

Les trois composantes de la quantité de mouvement et la quantité

$$\frac{m_0 c}{\sqrt{1-\beta^2}} \left(= \frac{W}{c} \text{ s'il n'y a pas de champ} \right)$$

constituent les 4 composantes d'un vecteur d'espace-temps.

Les formules précédentes doivent être modifiées dans le cas très important d'un corpuscule de charge électrique ε se déplaçant dans un champ électromagnétique. On sait que le champ électrique $\vec{h}$ et le champ magnétique $\vec{H}$ peuvent être définis à l'aide d'un potentiel scalaire $V(x, y, z, t)$ et d'un potentiel vecteur $\vec{A}(x, y, z, t)$ par les relations :

$$\vec{H} = \text{rot}\,\vec{A}\,; \quad \vec{h} = -\,\text{grad}\,V - \frac{1}{c}\frac{\partial \vec{A}}{\partial t} \tag{8}$$

La force qui s'exerce dans ce champ sur le corpuscule de charge ε animé de la vitesse $\vec{v}$ est donnée par la formule de Lorentz :

$$\vec{f} = \varepsilon \left(\vec{h} + \frac{1}{c} [\vec{v} \cdot \vec{H}] \right) \tag{9}$$

L'énergie du corpuscule est alors :

$$W = \frac{m_0 c^2}{\sqrt{1-\beta^2}} + \varepsilon V(x, y, z, t) \tag{10}$$

et sa quantité de mouvement est :

$$\vec{p} = \frac{m_0 \vec{v}}{\sqrt{1-\beta^2}} + \varepsilon \frac{\vec{A}(x, y, z, t)}{c} \tag{11}$$

Le fait nouveau ici est que la quantité de mouvement, tout comme l'énergie, contient, en plus du terme dépendant de la vitesse, un terme dépendant du champ : en général la quantité de mouvement $\vec{p}$ n'est plus dirigée suivant la vitesse. Les composantes de $\vec{p}$ et la quantité $\frac{W}{c}$ forment encore les composantes d'un vecteur d'espace-temps.

Les équations (10) et (11) nous fournissent la relation :

$$\frac{1}{c^2}(W - \varepsilon V)^2 - \sum_{x,y,z}\left(p_x - \frac{\varepsilon A_x}{c}\right)^2 - m_0^2 c^2 = 0 \tag{12}$$

qui va jouer un rôle essentiel pour obtenir une équation d'onde relativiste en Mécanique ondulatoire.

Si nous résolvons (12) par rapport à W, nous obtenons :

$$W = c\sqrt{m_0^2 c^2 + \sum_{x,y,z}\left(p_x - \frac{\varepsilon A_x}{c}\right)^2} + \varepsilon V \tag{13}$$

Le second membre de cette équation peut être désigné par $H(x, y, z, p_x, p_y, p_z, t)$ et est la fonction Hamiltonienne relativiste. Ce n'est pas une fonction rationnelle.

2. *La Mécanique ondulatoire relativiste.* — Pour obtenir l'équation d'onde de la Mécanique ondulatoire relativiste, il paraît tout naturel d'opérer comme au paragraphe 2 du chapitre V en partant non plus des formules de la Dynamique classique, mais de celles de la Dynamique relativiste. Malheureusement il se présente tout de suite une difficulté : la fonction Hamiltonienne définie par le second membre de (13) n'étant pas rationnelle, l'expression que l'on obtient en

y remplaçant p_x par $-\frac{h}{2\pi i}\frac{\partial}{\partial x}$ etc. n'est elle-même pas rationnelle en $\frac{\partial}{\partial x}$ etc. et ne représente pas un opérateur bien défini. On ne peut donc pas appliquer littéralement la méthode du paragraphe 2, chapitre V, c'est-à-dire poser comme équation d'onde : $H(\Psi) = \frac{h}{2\pi i}\frac{\partial \Psi}{\partial t}$.

Il y a cependant un moyen indirect de tourner la difficulté en utilisant l'équation (12) au lieu de (13). Pour cela on remarque que déjà dans la méthode non relativiste du chapitre V, quand on passe de l'équation classique $H = E$ à l'équation d'onde $H(\Psi) = \frac{h}{2\pi i}\frac{\partial \Psi}{\partial t}$, on remplace en somme l'énergie E par l'opérateur $\frac{h}{2\pi i}\frac{\partial}{\partial t}$. Il paraît donc naturel de remplacer ici W par cet opérateur; et cela est d'autant plus logique qu'au point de vue relativiste l'énergie et la quantité de mouvement formant un vecteur d'espace-temps, si l'on remplace chaque p_i par $-\frac{h}{2\pi i}\frac{\partial}{\partial q_i}$, W doit être remplacé par $\frac{h}{2\pi i}\frac{\partial}{\partial t}$ (1).

En faisant ces substitutions dans le premier membre de (12), on obtient

$$\frac{1}{c^2}\left(\frac{h}{2\pi i}\frac{\partial}{\partial t} - \varepsilon V\right)^2 - \sum_{x,y,z}\left(\frac{h}{2\pi i}\frac{\partial}{\partial x} + \frac{\varepsilon}{c}A_x\right)^2 - m_0^2 c^2 \qquad (14)$$

c'est-à-dire un opérateur rationnel.

En appliquant l'opérateur (14) à la fonction Ψ et en égalant à zéro, on trouve :

---

(1) La différence de signe s'explique en remarquant que

$$\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}, \frac{\partial}{\partial t}$$

sont des composantes covariantes tandis que $p_x\, p_y\, p_z$ W sont les composantes contrevariantes de l'Impulsion d'Univers.

$$\frac{1}{c^2}\left(\frac{h}{2\pi i}\frac{\partial}{\partial t} - \varepsilon V\right)^2 \Psi - \sum_{x,y,z}\left(\frac{h}{2\pi i}\frac{\partial}{\partial x} + \frac{\varepsilon}{c} A_x\right)^2 \Psi = m_0^2 c^2 \Psi \tag{15}$$

et cette équation peut être regardée comme l'extension relativiste naturelle de l'équation d'onde de la Mécanique ondulatoire primitive.

Si nous développons l'équation (15) en tenant compte de la relation de Lorentz entre les potentiels :

$$\frac{1}{c}\frac{\partial V}{\partial t} + \operatorname{div} \vec{A} = 0 \tag{16}$$

nous obtenons :

$$\frac{1}{c^2}\frac{\partial^2 \Psi}{\partial t^2} - \Delta\Psi - \frac{4\pi i}{h}\frac{\varepsilon V}{c^2}\frac{\partial \Psi}{\partial t} - \frac{4\pi i}{h}\sum_{x,y,z}\frac{\varepsilon A_x}{c}\frac{\partial \Psi}{\partial x} + \frac{4\pi^2}{h^2}\left[m_0^2 c^2 + \frac{\varepsilon^2}{c^2}(V^2 - A^2)\right]\Psi = 0 . \tag{17}$$

Pour des raisons que nous aurons à exposer plus tard, Dirac considère cette équation de propagation relativiste comme insuffisante. Nous voyons d'ailleurs tout de suite qu'elle diffère beaucoup de l'équation non relativiste à un point de vue important : c'est une équation différentielle du second ordre par rapport au temps au lieu d'être une équation différentielle du premier ordre.

Dans le cas important où le champ électromagnétique est nul ($V = \vec{A} = 0$), l'équation (17) s'écrit :

$$\Delta\Psi - \frac{1}{c^2}\frac{\partial^2 \Psi}{\partial t^2} = \frac{4\pi^2}{h^2} m_0^2 c^2 \Psi \tag{18}$$

et elle admet comme solution particulière l'onde monochromatique plane :

$$\Psi = C\, e^{\frac{2\pi i}{h}[Wt - p_x x - p_y y - p_z z]} \tag{19}$$

car on a alors $\frac{W^2}{c^2} - p_x^2 - p_y^2 - p_z^2 = m_0^2 c^2$ d'après (12).

L'onde (19) possède la fréquence $\nu = \frac{W}{h}$ et la longueur d'onde

$\lambda = \frac{h}{p}$; elle correspond au mouvement rectiligne uniforme d'un corpuscule dont l'énergie et la quantité de mouvement sont parfaitement déterminées alors que sa position est entièrement indéterminée. Si l'on compare la formule (19) avec la formule (7) du chapitre V, on verra que la seule différence essentielle introduite ici par la Relativité est la substitution de l'énergie complète W à l'énergie $E = W - m_0 c^2$ de la Mécanique classique.

3. ***Densité et courant de probabilité correspondant à l'équation*** (17). — L'application des principes généraux énoncés aux chapitres V et VI à l'équation d'onde (17) soulève de grandes difficultés. En particulier cela n'a plus un sens clair de dire que la fonction d'onde Ψ doit être normée car on ne peut plus démontrer avec l'équation (17) que, si Ψ est normée à un instant donné, elle reste ensuite toujours normée. Cette difficulté se rattache essentiellement, comme nous le verrons au chapitre X, à ce que l'équation (17) étant du second ordre par rapport au temps, sa solution n'est pas déterminée quand on connaît seulement la forme initiale de la fonction Ψ.

Il est néanmoins encore possible avec l'équation (17) de considérer un fluide de probabilité qui se conserve au cours du temps en vertu même de l'équation de propagation, mais la densité de ce fluide s'exprime en fonction non seulement de Ψ, mais aussi de $\frac{\partial \Psi}{\partial t}$. Posons en effet :

$$\rho = -\frac{h}{4\pi i m_0 c^2}\left(\Psi \frac{\partial \Psi^*}{\partial t} - \Psi^* \frac{\partial \Psi}{\partial t}\right) - \frac{\varepsilon}{m_0 c^2} V \Psi\Psi^* \quad (20)$$

$$\rho \vec{u} = \frac{h}{4\pi i m_0}\left(\Psi \operatorname{grad} \Psi^* - \Psi^* \operatorname{grad} \Psi\right) - \frac{\varepsilon}{m_0 c} \vec{A} \Psi\Psi^* \quad (20')$$

équations qui définissent la répartition et le mouvement d'un fluide fictif de probabilité car ρ et $\rho\vec{u}$ sont réels :

Ecrivons maintenant l'équation conjuguée de (17)

$$\frac{1}{c^2}\frac{\partial^2 \Psi^*}{\partial t^2} - \Delta \Psi^* + \frac{4\pi i}{h}\frac{\varepsilon}{c^2} V \frac{\partial \Psi}{\partial t}$$
$$+ \frac{4\pi i}{h}\frac{\varepsilon}{c}\sum_{x,y,z} A_x \frac{\partial \Psi}{\partial x} + \frac{4\pi^2}{h^2}\left[m_0{}^2 c^2 + \frac{\varepsilon^2}{c^2}(V^2 - A^2)\right]\Psi = 0 . \tag{17*}$$

Multiplions (17) par Ψ^* et retranchons (17*) multiplié par Ψ. On trouve d'une part les termes :

$$\frac{1}{c^2}\left(\Psi^* \frac{\partial^2 \Psi}{\partial t^2} - \Psi \frac{\partial^2 \Psi^*}{\partial t^2}\right) - \frac{4\pi i}{h}\frac{\varepsilon}{c^2} V \left(\Psi \frac{\partial \Psi^*}{\partial t} + \Psi^* \frac{\partial \Psi}{\partial t}\right)$$

qu'on peut écrire :

$$\frac{\partial}{\partial t}\left[\frac{1}{c^2}\left(\Psi^* \frac{\partial \Psi}{\partial t} - \Psi \frac{\partial \Psi^*}{\partial t}\right) - \frac{4\pi i}{h}\frac{\varepsilon}{c^2} V \Psi\Psi^*\right]$$
$$+ \frac{4\pi i}{h}\frac{\varepsilon}{c^2}\Psi\Psi^* \frac{\partial V}{\partial t}$$

et d'autre part, on a les termes :

$$-\Psi^* \Delta\Psi + \Psi\Delta\Psi^* - \frac{4\pi i}{h}\frac{\varepsilon}{c}\sum_{x,y,z}\left(A_x \Psi^* \frac{\partial \Psi}{\partial x} + A_x \Psi \frac{\partial \Psi^*}{\partial x}\right)$$

qu'on peut écrire :

$$\sum_{x,y,z}\left\{ \frac{\partial}{\partial x}\left[\left(\Psi \frac{\partial \Psi^*}{\partial x} - \Psi^* \frac{\partial \Psi}{\partial x}\right) - \frac{4\pi i}{h}\frac{\varepsilon}{c} A_x \Psi\Psi^*\right]\right.$$
$$\left. + \frac{4\pi i}{h}\frac{\varepsilon}{c}\Psi\Psi^* \frac{\partial A^*}{\partial x}\right\}$$

Au total, nous obtenons l'équation :

$$\frac{\partial}{\partial t}\left[\frac{1}{c^2}\left(\Psi^* \frac{\partial \Psi}{\partial t} - \Psi \frac{\partial^* \Psi}{\partial t}\right) - \frac{4\pi i}{h}\frac{\varepsilon}{c^2} V \Psi\Psi^*\right]$$
$$+ \operatorname{div}\left[\Psi \operatorname{grad} \Psi^* - \Psi^* \operatorname{grad} \Psi - \frac{4\pi i}{h}\frac{\varepsilon}{c}\vec{A}\Psi\Psi^*\right]$$
$$+ \frac{4\pi i}{h}\frac{\varepsilon}{c}\Psi\Psi^*\left(\frac{1}{c}\frac{\partial V}{\partial t} + \operatorname{div}\vec{A}\right) = 0 . \tag{21}$$

Le dernier terme est nul en raison de la relation de Lorentz (16). Multiplions par $\dfrac{h}{4\pi i m_0}$ et tenons compte de (20) et (20'); il vient :

$$\frac{\partial \rho}{\partial t} + \operatorname{div} (\rho \overrightarrow{u}) = 0 \qquad (22)$$

c'est-à-dire l'équation qui exprime la conservation du fluide (¹).

A la répartition et au mouvement du fluide de probabilité associé à un corpuscule de charge électrique ε, nous pouvons faire correspondre une densité électrique $\rho\,\varepsilon$ et une densité de courant $\rho\,\varepsilon\,\overrightarrow{u}$ comme nous l'avons fait à la fin du chapitre précédent dans le cas de non-relativité.

On se rendra compte des difficultés qu'il y a à concilier les principes généraux admis au chapitre V avec la forme relativiste (17) de l'équation des ondes en se reportant à la façon dont nous avons justifié le principe des interférences au paragraphe 5 de ce chapitre V. Du raisonnement exposé à cet endroit, il résulte que la probabilité pour qu'une mesure faite à l'instant t permette de localiser le corpuscule dans un élément $d\tau$ de l'espace est $\Psi\Psi^* d\tau$; ce résultat est obtenu sans faire aucune hypothèse sur l'équation de propagation à laquelle obéit la fonction d'onde Ψ et il devrait donc être encore valable ici. Cependant la densité ρ du fluide de probabilité donnée par la relation (20), densité dont l'expression est imposée par la nécessité de satisfaire la condition de conservation, ne se réduit pas à $\Psi\Psi^*$. Il y a donc contradiction entre les principes généraux du chapitre V et la forme relativiste (17) de l'équation de propagation. C'est le souci d'éviter une telle contradiction qui a conduit M. Dirac à fonder la mécanique ondulatoire relativiste du corpuscule sur une base différente de l'équation (17).

(¹) Il semblerait qu'on aurait pu définir ρ et $\rho\,\overrightarrow{u}$ en prenant les seconds membres de (20) et (20′) multipliés par une constante réelle quelconque, mais les définitions (20) et (20′) s'imposent si l'on veut retrouver à l'approximation newtonienne l'expression $\rho = \Psi\Psi^*$ comme on s'en rendra compte aisément.

CHAPITRE VIII

Succès et échecs de la Mécanique ondulatoire à une fonction d'onde

1. *Calcul des énergies quantifiées; exemple de l'atome d'hydrogène*

PAR application de ses principes généraux, la Mécanique ondulatoire (non relativiste) détermine l'énergie des états stationnaires pour les systèmes quantifiés en calculant les valeurs propres de l'opérateur Hamiltonien correspondant. Cette nouvelle méthode de quantification inaugurée par les célèbres travaux de M. Schrödinger [1] a conduit à retrouver dans certains cas le résultat de l'ancienne théorie des quanta et dans d'autres cas elle a corrigé les résultats anciens dans un sens plus conforme à l'expérience (oscillateur linéaire).

Nous rappellerons ici seulement la quantification de l'atome d'hydrogène et nous nous servirons comme Schrödinger de la théorie non relativiste.

Nous considérons un électron de masse m et de charge $-e$ dans le champ produit par un noyau fixe de charge $+e$.

Ecrivons l'équation :

$$H(a) = E \cdot a \tag{1}$$

(1) Voir E. SCHRÖDINGER, *Abhandlungen zur Wellenmechanik*, J. A. Barth, Leipzig.

en prenant pour H l'opérateur $\frac{1}{2m}\,[p_x^2+p_y^2+p_z^2]-\frac{e^2}{r}$.

Nous obtenons : $\Delta a+\frac{8\pi^2 m}{h^2}\left[E+\frac{e^2}{r}\right]a=0.$ (2)

Prenons des coordonnées polaires r, θ, φ autour du noyau et posons :

$$a(r,\theta,\varphi)=R(r)\,Y(\theta,\varphi) \qquad (3)$$

En tenant compte de la forme du Laplacien en coordonnées polaires, nous avons :

$$Y\frac{d^2 R}{dr^2}+2\frac{Y}{r}\frac{dR}{dr}+\frac{1}{r^2\sin^2\theta}\frac{\partial^2 Y}{\partial\varphi^2}+$$
$$\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right)+\frac{8\pi^2 m}{h^2}\left[E+\frac{e^2}{r}\right]RY=0 \qquad (4)$$

ou encore :

$$r^2\left[\frac{1}{R}\frac{d^2 R}{dr^2}+\frac{2}{r}\frac{1}{R}\frac{dR}{dr}+\frac{8\pi^2 m}{h}\left(E+\frac{e^2}{r}\right)\right]=$$
$$-\frac{1}{Y}\left[\frac{1}{\sin^2\theta}\frac{\partial^2 Y}{\partial\varphi^2}+\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right)\right] \qquad (5)$$

Les deux membres de (5) dépendant l'un du rayon vecteur seulement, l'autre des angles polaires seulement, doivent être égaux à une même constante λ et l'on a :

$$\frac{1}{\sin^2\theta}\frac{\partial^2 Y}{\partial\varphi^2}+\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right)+\lambda Y=0 \qquad (6)$$

On démontre que l'équation (6) n'a de solutions finies, uniformes et continues sur la sphère du rayon un que si l'on a :

$$\lambda=l(l+1) \quad \text{avec} \quad l=0,\ 1,\ 2\ldots \qquad (7)$$

Les valeurs (7) de λ sont les valeurs propres de l'équation (6). A la valeur propre définie par une certaine valeur entière de l correspondent les $(2l+1)$ fonctions propres (non normées) :

$$Y_l^m(\theta,\varphi)=e^{im\varphi}P_l^m(\cos\theta)=e^{im\varphi}\sin^m\theta\,\frac{d^{l+m}}{(d\cos\theta)^{l+m}}\cdot(1-\cos^2\theta)^l \qquad (8)$$

Les fonctions Y_l^m sont les fonctions sphériques de Laplace. Elles forment un système complet pour les variables φ et θ, ce qui justifie *a posteriori* la décomposition (3).

Ceci posé, il résulte encore de (5) que R doit satisfaire à l'équation :

$$\frac{d^2 R}{dr^2} + \frac{2}{r}\frac{dR}{dr} + R\left(A + \frac{B}{r} + \frac{C}{r^2}\right) = 0 \tag{9}$$

avec :

$$A = \frac{8\pi^2 m E}{h^2}\,;\ B = \frac{8\pi^2 m}{h^2} e^2\,;\ C = l(l+1) \tag{10}$$

Schrödinger a démontré que toutes les valeurs positives de E sont valeurs propres de (9) et forment par suite un spectre continu. Mais ces valeurs propres positives de l'énergie correspondent, comme en Mécanique classique, à des mouvements libres de l'électron hors de l'atome et ne nous intéressent pas ici.

Pour trouver les valeurs propres négatives, introduisons la variable réelle :

$$\rho = 2\sqrt{-A}\, r = \frac{4\pi}{h}\sqrt{-2mE}\,.\, r \tag{11}$$

Il est visible sur l'équation (9) que, pour r très grand, R a la forme asymptotique $e^{-\frac{\rho}{2}}$, la solution $e^{+\frac{\rho}{2}}$ devant être écartée car elle est infinie à l'infini. Posons donc :

$$R = e^{-\frac{\rho}{2}} v(\rho) \tag{12}$$

En substituant (11) et (12) dans (9), nous trouvons aisément :

$$\frac{d^2 v}{d\rho^2} + \left(\frac{2}{\rho} - 1\right)\frac{dv}{d\rho} + \left[\left(\frac{B}{2\sqrt{-A}} - 1\right)\frac{1}{\rho} - \frac{l(l+1)}{\rho^2}\right] v = 0 \tag{13}$$

Ceci est une équation différentielle linéaire admettant à distance finie le seul point singulier $\rho = 0$. La théorie des équations linéaires permet alors de voir aisément que l'équation (13) admet une seule solution régulière au voisinage du point $\rho = 0$ et finie en ce point et que cette solution est de la forme :

$$v(\rho) = \sum_\nu a_\nu \rho^{l+\nu} \tag{14}$$

En substituant (14) dans (13), nous trouvons la relation de récurrence :

$$[(\nu+l+1)(\nu+1)+2(\nu+l+1)-l(l+1)]\,a_{\nu+1} = \left[\nu+l+1-\frac{B}{2\sqrt{-A}}\right]a_\nu \qquad (15)$$

La fonction R(ρ) sera nulle à l'infini si tous les a_ν sont nuls à partir d'un certain a_p ; d'après la formule (15), il faut pour cela que l'on ait :

$$\frac{B}{2\sqrt{-A}} = p+l+1 = n \qquad (n \text{ entier} \geqslant 1) \qquad (16)$$

D'où d'après les valeurs (10) :

$$E_n = -\frac{2\pi^2 m e^4}{n^2 h^2} \qquad (17)$$

La Mécanique ondulatoire non relativiste donne donc la formule de Bohr.

Remarquons qu'il y a ici une dégénérescence accentuée car à une valeur propre E_n, c'est-à-dire à une valeur donnée de n, correspondent d'après (16) n valeurs possibles de l savoir 0, 1, ..., $n-1$ et à chaque valeur de l correspondent $2l+1$ fonctions sphériques Y_l^m. Donc à une valeur propre E_n, correspondent des fonctions propres $u = RY$ en nombre égal à $\sum_0^{n-1}{}_l (2l+1) = 2\frac{n(n-1)}{2} + n = n^2$. Toutes les valeurs propres sont donc multiples sauf celle qui correspond à $n=1$.

Le nombre quantique l correspond au nombre $k-1$ de l'ancienne théorie des quanta ; il peut prendre les valeurs 0, 1 ..., $n-1$ alors que le nombre k de l'ancienne théorie prenait les valeurs 1, 2 ... n. Le nombre m dépend du choix de l'axe polaire qui est entièrement arbitraire en l'absence de champ extérieur. Il est aisé de vérifier que les moments de rotation $M_x M_y M_z$ sont des intégrales premières comme cela doit être à cause de la symétrie sphérique du champ Coulombien.

Notons enfin qu'en reprenant le calcul pour un atome de charge centrale $+Ne$ ionisé $(N-1)$ fois, on trouve

$$E_n = -\frac{2\pi^2 m e^4 N^2}{n^2 h^2} \tag{18}$$

d'où l'on peut tirer comme dans l'ancienne théorie une justification approchée de la loi de Moseley.

2. ***La structure fine et la Mécanique ondulatoire relativiste.*** — En calculant les valeurs possibles de l'énergie pour l'atome d'hydrogène, nous avons retrouvé le résultat simple de Bohr. Il est tout naturel de nous demander maintenant si, en employant la forme relativiste de la Mécanique ondulatoire développée au chapitre précédent, nous pourrons retrouver la structure fine de Sommerfeld. Mais il se présente aussitôt une difficulté, nous ne pouvons chercher à déterminer les valeurs propres de l'opérateur Hamiltonien parce que cet opérateur dans la théorie du dernier chapitre n'est pas bien déterminé, nous l'avons vu. Néanmoins il y a une manière assez naturelle (bien que mal en accord avec les principes généraux) de tourner la difficulté. En effet, chercher les valeurs propres de l'énergie revient à chercher les fréquences propres de l'équation des ondes. Prenons donc l'équation d'onde relativiste (17) du dernier chapitre et supposons que Ψ soit une onde monochromatique ne dépendant du temps que par le facteur $e^{\frac{2\pi i}{h}Wt}$. On trouve alors en notant que dans l'atome d'hydrogène $\vec{A}=0$:

$$\Delta\Psi + \frac{4\pi^2}{h^2c^2}\left[(W-\varepsilon V)^2 - m_0^2c^4\right]\Psi = 0 \tag{19}$$

Dans le cas de l'atome d'hydrogène, $\varepsilon = -e$ et $V = \frac{e}{r}$; donc le crochet de (19) s'écrit :

$$\left(m_0c^2 + E + \frac{e^2}{r}\right)^2 - m_0^2c^4 = E^2 + 2m_0c^2E + \frac{2e^2}{r}(m_0c^2 + E) + \frac{e^4}{r^2} \tag{20}$$

et comme par hypothèse $\Psi = a(r, \theta, \varphi)\, e^{\frac{2\pi}{h} Wt}$, on a :

$$\Delta a + \frac{4\pi^2}{h^2 c^2}\left[E^2 + 2m_0 c^2 E + \frac{2e^2}{r}(m_0 c^2 + E) + \frac{e^4}{r^2}\right] a = 0 \qquad (21)$$

En posant encore ici la décomposition (3), on voit aisément que $Y(\theta, \varphi)$ doit être une fonction sphérique de Laplace et que $R(r)$ doit obéir à l'équation :

$$\frac{d^2 R}{dr^2} + \frac{2}{r}\frac{dR}{dr} + \left[A + \frac{B}{r} + \frac{C}{r^2}\right] R = 0 \qquad (22)$$

avec les notations :

$$A = \frac{8\pi^2 m_0}{h^2} E\left(1 + \frac{E}{2m_0 c^2}\right) ;\ B = \frac{8\pi^2 m_0}{h^2} e^2 \left(1 + \frac{E}{m_0 c^2}\right);$$

$$C = -l(l+1) + \frac{4\pi^2 e^4}{h^2 c^2}. \qquad (23)$$

L'expression de C peut aussi s'écrire :

$$C = -l(l+1) + \alpha^2 \qquad (23')$$

où $\alpha = \frac{2\pi e^2}{h^2 e^2}$ est la constante de structure fine de Sommerfeld.

De l'équation (22) nous déduisons que R a la forme asymptotique $e^{-\frac{\rho}{2}}$ avec $\rho = 2\sqrt{-A} \,.\, r$ et en posant :

$$R(\rho) = e^{-\frac{\rho}{2}} v(\rho) \qquad (24)$$

nous obtenons par substitution dans (22) :

$$\frac{d^2 v}{d\rho^2} + \left(\frac{2}{\rho} - 1\right)\frac{dv}{d\rho} + \left[\left(\frac{B}{2\sqrt{-A}} - 1\right)\frac{1}{\rho} + \frac{C}{\rho^2}\right] v = 0 \qquad (25)$$

C'est une équation différentielle linéaire admettant comme seule singularité à distance finie le point $\rho = 0$. D'après la théorie générale des équations linéaires, il existe, au plus, deux solutions de (25) qui sont régulières au voisinage de $\rho = 0$ et de la forme :

$$v(\rho) = \rho^\gamma \sum_\nu a_\nu \rho^\nu \qquad (a_0 \neq 0) \qquad (26)$$

L'exposant γ (non nécessairement entier) est donné par l'équation déterminante :

$$\gamma(\gamma-1)+2\gamma+C=\gamma(\gamma+1)+\alpha^2-l(l+1)=0 \quad (27)$$

qu'on peut écrire aussi bien :

$$\left(\gamma+\frac{1}{2}\right)^2=\left(l+\frac{1}{2}\right)^2-\alpha^2 \quad (28)$$

On a donc les deux valeurs de γ :

$$\gamma=-\frac{1}{2}\pm\sqrt{\left(l+\frac{1}{2}\right)^2-\alpha^2} \quad (29)$$

On écarte la solution correspondant au signe — en remarquant que la fonction $v(\rho)$ correspondante aurait un pôle pour $\rho=0$. Nous garderons donc la formule (29) avec le signe $+$. Il convient de remarquer que même avec le signe $+$ elle soulève une petite difficulté si $l=0$ car elle donne alors une valeur négative très petite de γ et la fonction $v(\rho)$ est infinie (d'un ordre très petit, il est vrai) en $\rho=0$; nous conviendrons de passer sur cette difficulté que nous retrouverons dans la théorie de Dirac parce que la fonction $v(\rho)$ bien qu'infinie en $\rho=0$, est cependant de carré sommable.

Si nous introduisons la forme (26) dans (25), nous obtenons la relation de récurrence :

$$[(\nu+\gamma+1)(\nu+\gamma)+2(\nu+\gamma+1)+C]a_{\nu+1}=$$
$$=\left[\nu+\gamma+1-\frac{B}{2\sqrt{-A}}\right]a_\nu \quad (30)$$

La fonction R sera nulle à l'infini si tous les a_ν supérieurs à un certain a_p sont nuls, ce qui arrivera si l'on a :

$$p+\gamma+1=\frac{B}{2\sqrt{-A}} \quad (31)$$

En substituant dans (31) les valeurs de A, de B et de γ, on trouve :

$$\frac{\alpha\left(1+\frac{E}{m_0 c^2}\right)}{\sqrt{-\frac{2E}{m_0 c^2}\left(1+\frac{E}{2m_0 c^2}\right)}} = \left(p+\frac{1}{2}\right)+\sqrt{\left(l+\frac{1}{2}\right)^2-\alpha^2} \tag{32}$$

d'où :

$$\frac{\alpha^2}{\left[\left(p+\frac{1}{2}\right)+\sqrt{\left(l+\frac{1}{2}\right)^2-\alpha^2}\right]^2} = \frac{-\frac{2E}{m_0 c^2}\left(1+\frac{E}{2m_0 c^2}\right)}{\left(1+\frac{E}{m_0 c^2}\right)^2} \tag{33}$$

En ajoutant l'unité aux deux membres et en prenant les inverses, on obtient aisément :

$$1+\frac{E}{m_0 c^2} = \left[1+\frac{\alpha^2}{\left[\left(p+\frac{1}{2}\right)+\sqrt{\left(l+\frac{1}{2}\right)^2-\alpha^2}\right]}\right]^{-\frac{1}{2}} \tag{34}$$

C'est une formule très analogue à la formule ancienne de Sommerfeld [formule (38) du chapitre I] mais où les nombres entiers n_2 et n_1 sont respectivement remplacés par

$$p+\frac{1}{2} \text{ et } l+\frac{1}{2}.$$

Si nous définissons un nombre quantique total n par la relation :

$$n=\left(p+\frac{1}{2}\right)+\left(l+\frac{1}{2}\right)=p+l+1 \tag{35}$$

et si nous développons (34) suivant les puissances de α^2 en négligeant les termes d'ordre supérieur au second, nous obtenons la formule approchée :

$$E_{nl}=-\frac{Rh}{n^2}\left[1+\frac{\alpha^2}{n^2}\left(\frac{n}{l+\frac{1}{2}}-\frac{3}{4}\right)\right] \tag{36}$$

R=C^te de Rydberg. La formule (36) est à comparer à la formule (41) du chapitre I.

Quand on reprend le même calcul pour un atome de rang N ionisé (N — 1) fois, on trouve au lieu de (36) :

$$E_{nl} = -\frac{RhN^2}{n^2}\left[1 + \frac{\alpha^2 N^2}{n^2}\left(\frac{n}{l+\frac{1}{2}} - \frac{3}{4}\right)\right] \quad (37)$$

analogue à la formule (44) du chapitre I.

Dans le cas des Rayons X, en se bornant à considérer grossièrement les électrons intérieurs comme formant écran électrostatique, on obtient :

$$E_{nl} = -\frac{Rh\,(N - z_{nl})}{n^2}\left[1 + \frac{\alpha^2\,(N - z_{nl})}{n^2}\left(\frac{n}{l+\frac{1}{2}} - \frac{3}{4}\right)\right] \quad (38)$$

et cette formule donnant la structure fine des termes spectraux X est à comparer avec l'ancienne formule (9) du chapitre III.

Malheureusement l'accord des anciennes formules de Sommerfeld avec l'expérience ne se retrouve pas ici. Prenons par exemple les doublets de la série de Balmer. Une raie de la série de Balmer, dans le schéma primitif de Bohr, est engendrée lors du passage de l'atome H d'un état initial d'énergie E_i à un état final caractérisé par $n = 2$. Pour Sommerfeld, au nombre quantique $n = 2$ correspondent deux valeurs possibles du nombre quantique azimutal $k = 1$ et $k = 2$; d'où une structure fine de chaque raie qui est en réalité un doublet, l'écartement du doublet en fréquence devant être d'après la formule (40) du chapitre I égal à :

$$\delta\nu = \frac{1}{h}\,(E_{22} - E_{21}) = \frac{R\alpha^2}{2^4}\left(\frac{2}{1} - \frac{2}{2}\right) = \frac{R\alpha^2}{16} = 0{,}36\ \text{cm}^{-1}. \quad (39)$$

Ce nombre quoiqu'un peu fort est en assez bon accord avec l'expérience. Avec notre nouvelle formule (36), nous devons attribuer aux deux niveaux responsables des doublets de Balmer les nombres quantiques $n = 2$, $l = 0$ et $n = 2$, $l = 1$ et nous trouvons :

$$\delta\nu = \frac{R\alpha^2}{2^4}\left(\frac{2}{\frac{1}{2}} - \frac{2}{\frac{3}{2}}\right) = \frac{R\alpha^2}{16}\cdot\frac{8}{3}. \qquad (40)$$

Nous trouvons donc les $\frac{8}{3}$ de l'ancien nombre qui était déjà plutôt trop grand !

De même pour les Rayons X, on trouverait par la formule (38) un nombre trop grand pour l'écartement des doublets dits de Sommerfeld.

De plus il reste toujours ici la difficulté déjà signalée : il est certain, du moins pour les Rayons X, que les niveaux responsables de la structure fine de Sommerfeld ont le même nombre quantique k (c'est-à-dire le même nombre $l = k - 1$) et diffèrent par un nombre quantique j que la théorie précédente ignore totalement.

Ainsi donc la Mécanique ondulatoire relativiste développée au précédent chapitre se montre insuffisante puisque, sans parvenir à résoudre les anciennes difficultés, elle en introduit au contraire de nouvelles.

3. *La Mécanique ondulatoire à une fonction d'onde et l'effet Zeeman.* — Nous allons voir maintenant que la Mécanique ondulatoire primitive (à une fonction d'onde) ne peut pas rendre compte de l'effet Zeeman anomal.

Considérons un atome plongé dans un champ magnétique uniforme $\vec{H}$. Prenons la direction du champ pour axe des z : nous pouvons alors écrire le potentiel vecteur dont dérive le champ $\vec{H}$ sous la forme.

$$A_x = -\frac{Hy}{2} \qquad A_y = \frac{Hx}{2} \qquad A_z = 0 \qquad (41)$$

car la relation $\vec{H} = \text{rot}\,\vec{A}$ donne alors $H_x = H_y = 0$ et $H_z = H$. L'équation d'onde relativiste (17) du dernier chapitre s'écrit alors ($\varepsilon = -e$) :

$$\Delta\Psi - \frac{1}{c^2}\frac{\partial^2\Psi}{\partial t^2} + \frac{4\pi i eV}{h}\frac{\partial\Psi}{c^2}\frac{\partial\Psi}{\partial t} + \sum_{xyz}\frac{4\pi i}{h}\frac{e}{c}A_x\frac{\partial\Psi}{\partial x}$$
$$= \frac{4\pi^2}{h^2}\left[m_0{}^2c^2 - \frac{e^2}{c^2}(V^2 - A^2)\right]\Psi = 0\,. \qquad (42)$$

Nous allons démontrer le théorème suivant qui n'est autre que la transposition en Mécanique ondulatoire du théorème classique de Larmor rappelé au chapitre IV, parag. 2 :

Théorème : Lorsque le champ H est assez faible et que l'on néglige les corrections de relativité, l'équation d'onde de l'atome exprimée dans un système de référence qui tourne autour de la direction du champ H avec la vitesse angulaire :

$$o = \frac{1}{2}\frac{eH}{m_0c} \qquad (43)$$

est la même que si le système de référence ne tournait pas et si le champ H n'existait pas.

D'après nos hypothèses, nous supposerons H petit et H^2 négligeable ; de même nous supposerons $\eta = \frac{E}{m_0c^2}$ petit (approximation Newtonienne) et η^2 négligeable. Enfin nous négligerons aussi le produit ηH.

Nous prenons toujours un système d'axes *oxyz* directs. L'axe *oz* servant d'axe cylindrique, définissons les coordonnées cylindriques z, ρ, φ par les formules usuelles :

$$x = \rho\cos\varphi \qquad y = \rho\sin\varphi \qquad z = z \qquad (44)$$

On a alors grâce à (41) :

$$\frac{4\pi i}{h}\frac{e}{c}\sum_{xyz}A_x\frac{\partial\Psi}{\partial x} = \frac{2\pi i}{h}\frac{e}{c}H\left(y\frac{\partial\Psi}{\partial x} - x\frac{\partial\Psi}{\partial y}\right)$$
$$= -\frac{2\pi i}{h}\frac{e}{c}H\frac{\partial\Psi}{\partial\varphi} \qquad (45)$$

et l'équation (42) s'écrit en coordonnées cylindriques en négligeant A^2 :

$$\frac{\partial^2\Psi}{\partial z^2} + \frac{\partial^2\Psi}{\partial\rho^2} + \frac{1}{\rho}\frac{\partial\Psi}{\partial\rho} + \frac{1}{\rho^2}\frac{\partial^2\Psi}{\partial\varphi^2} - \frac{1}{c^2}\frac{\partial^2\Psi}{\partial t^2} + \frac{4\pi i}{h}\frac{eV}{c^2}\frac{\partial\Psi}{\partial t}$$
$$-\frac{2\pi i}{h}\frac{e}{c}H\frac{\partial\Psi}{\partial\varphi} = \frac{4\pi^2}{h^2}\left(m_0{}^2c^2 - \frac{e^2V^2}{c^2}\right)\Psi\,. \qquad (46)$$

Prenons maintenant un système de coordonnées cylindriques z' ρ' φ' qui tourne autour de oz avec la vitesse angulaire (43). On a :

$$z' = z \quad \rho' = \rho \quad \varphi' = \varphi - ot \quad t' = t \tag{47}$$

et par suite (en négligeant o^2 de l'ordre de H^2) :

$$\frac{\partial}{\partial z} = \frac{\partial}{\partial z'} ; \ \frac{\partial}{\partial \rho} = \frac{\partial}{\partial \rho'} ; \ \frac{\partial}{\partial \varphi} = \frac{\partial}{\partial \varphi'} ; \ \frac{\partial}{\partial t} = \frac{\partial}{\partial t'} - o \frac{\partial}{\partial \varphi} ;$$
$$\frac{\partial^2}{\partial t^2} = \frac{\partial^2}{\partial t'^2} - 2\, o \frac{\partial^2}{\partial \varphi' \, \partial t'} \tag{47'}$$

L'équation (46) s'écrit donc dans le système des variables primées :

$$\frac{\partial^2 \Psi}{\partial z'^2} + \frac{\partial^2 \Psi}{\partial \rho'^2} + \frac{1}{\rho'} \frac{\partial \Psi}{\partial \rho'} + \frac{1}{\rho'^2} \frac{\partial^2 \Psi}{\partial \varphi'^2} - \frac{1}{c^2} \frac{\partial^2 \Psi}{\partial t'^2} + \frac{4 \pi i e V}{h} \frac{\partial \Psi}{c^2 \, \partial t'}$$
$$+ \frac{2o}{c^2} \frac{\partial^2 \Psi}{\partial \varphi' \partial t'} - \frac{4 \pi i e V}{h \; c^2} o \frac{\partial \Psi}{\partial \varphi'} - \frac{2 \pi i}{h} \frac{e}{c} H \frac{\partial \Psi}{\partial \varphi'}$$
$$= \frac{4 \pi^2}{h^2} \left(m_0^2 c^2 - \frac{e^2 V^2}{c^2} \right) \Psi \tag{48}$$

Or on a en tenant compte de (43) et en négligeant $o\, \eta$:

$$\frac{2\, o}{c^2} \frac{\partial^2 \Psi}{\partial \varphi' \partial t'} = \frac{2\, o}{c^2} \frac{\partial}{\partial \varphi'} \left[\frac{2 \pi i}{h} m_0 c^2 (1 + \eta) \Psi \right] = \frac{2 \pi i}{h} \frac{e}{c} H \frac{\partial \Psi}{\partial \varphi'} \tag{49}$$

Le terme $\frac{2o}{c^2} \frac{\partial^2 \Psi}{\partial \varphi' \partial t'}$, compense donc à l'approximation admise le dernier terme dans le premier membre de (48). Comme à l'approximation Newtonienne l'énergie potentielle doit être regardée comme très petite devant l'énergie interne $m_0 c^2$, le terme $\frac{4 \pi i e V}{h \; c^2} o \frac{\partial \Psi}{\partial \varphi'}$ est négligeable devant le terme $\frac{4 \pi i e V}{h \; c^2} \frac{\partial \Psi}{\partial t}$.

Finalement l'équation (48) se réduit à :

$$\frac{\partial^2 \Psi}{\partial z'^2} + \frac{\partial^2 \Psi}{\partial \rho'^2} + \frac{1}{\rho'} \frac{\partial \Psi}{\partial \rho'} + \frac{1}{\rho'^2} \frac{\partial^2 \Psi}{\partial \varphi'^2} - \frac{1}{c^2} \frac{\partial^2 \Psi}{\partial t'^2}$$
$$+ \frac{4 \pi i e V}{h \; c^2} \frac{\partial \Psi}{\partial t'} = \frac{4 \pi^2}{h^2} \left(m_0^2 c^2 - \frac{e^2 V^2}{c^2} \right) \Psi \tag{50}$$

Cette équation est la même que si le système primé ne tournait pas et si le champ H n'existait pas : le théorème de Larmor est donc démontré.

Nous allons alors retrouver tout de suite l'effet Zeeman normal. En effet, en l'absence de champ magnétique extérieur, l'état stationnaire caractérisé par les deux nombres quantiques n et l et d'énergie E^{H}_{nl} correspond dans le système fixe z, ρ, φ à une fonction d'onde de la forme :

$$\Psi (z, \rho, \varphi, t) = F (\rho, z)\, \epsilon^{im\varphi}\, e^{\frac{2\pi i}{h}\left(m_0 c^2 + E^{o}_{nl}\right) t} \qquad (51)$$

m étant un entier positif ou négatif. En vertu du théorème démontré plus haut, la fonction Ψ aura, en présence du champ H, une expression de même forme dans le système z', ρ', φ'. On aura donc :

$$\Psi (z', \rho', \varphi', t') = F' (\rho', z')\, e^{im\varphi'}\, e^{\frac{2\pi i}{h}\left(m_0 c^2 + E^{o}{}_{nl}\right) t'} \qquad (52)$$

En vertu de (47), la fonction d'onde aura comme expression dans le système fixe :

$$\Psi (\rho, \varphi, z, t) = F (\rho, z)\, e^{im(\varphi - ot)}\, e^{\frac{2\pi i}{h}\left(m_0 c^2 + E^{o}{}_{nl}\right) t}$$

$$= F (\rho, z)\, e^{im\varphi}\, e^{\frac{2\pi i}{h}\left(m_0 c^2 + E^{o}{}_{nl} - \frac{mh}{2\pi} o\right) t} \qquad (53)$$

L'énergie E^{H}_{nl} de l'atome dans le système fixe en présence du champ H est donc :

$$E^{\text{H}}_{nl} = E^{o}_{nl} - m \frac{ho}{2\pi} = E^{o}_{nl} - m \frac{e h H}{4 \pi m_0 c}. \qquad (54)$$

Les termes spectraux sont donc modifiés par l'adjonction d'un multiple de $\dfrac{eH}{4\pi m_0 c}$ et la raie de fréquence

$$\nu = \frac{1}{h}\left(E^{o}_{n'l'} - E^{o}_{nl}\right)$$

a en présence du champ de fréquence

$$\nu + (m - m')\,\frac{e\mathrm{H}}{4\,\pi\, m_0\, c}.$$

Il suffit d'admettre la règle de sélection $\delta m = 0, \pm 1$ que nous justifierons plus loin pour retrouver l'effet Zeeman normal, mais il n'y a pas trace dans cette théorie des anomalies et du facteur de Landé. La Mécanique ondulatoire à une fonction d'onde ne donne donc rien de plus que la théorie classique ou que l'ancienne théorie des quanta.

4. *Les « règles de sélection » en Mécanique ondulatoire.* — En électromagnétisme classique, quand une distribution d'électricité est animée de mouvements accélérés, il y a émission d'une radiation électromagnétique. Si $\rho(x, y, z, t)$ est la densité de la distribution électrique, l'intensité du rayonnement émis est déterminée en première approximation par le moment électrique c'est-à-dire par le vecteur $\vec{m}$ de composantes rectangulaires données par :

$$m_x = \iiint \rho\, x\, dx\, dy\, dz\,; \quad m_y = \iiint \rho\, y\, dx\, dy\, dz\,;$$

$$m_z = \iiint \rho\, z\, dx\, dy\, dz\,. \tag{55}$$

Si l'on suppose les quantités (55) développées en séries de Fourier sous la forme :

$$m_x = \mathrm{C}^{\mathrm{t}} + \sum_i m_x^i \cos\left(2\,\pi\nu_i\, t + \varepsilon_i\right) \tag{56}$$

l'énergie émise par seconde en rayonnement de fréquence ν_i dont la vibration électrique est parallèle à l'axe des x est égale à $\dfrac{64\,\pi^4}{3\,c^3}\,\nu_i^4\,(m_x^i)^2$.

La connaissance du moment électrique d'une distribution permet donc de prévoir les fréquences, polarisations et intensités du rayonnement émis.

Avec les idées classiques, la radiation est émise d'une façon continue et l'énergie rayonnée est empruntée progressivement au mouvement de la distribution électrique, mouvement qui s'amortit en conséquence. Dans la théorie des quanta, les choses se présentent très différemment : la radiation est émise sous la forme d'un quantum $h\nu$ lors du passage d'un état quantifié à un autre. Pour évaluer l'intensité émise sous forme de radiation par une assemblée d'atomes émetteurs, il faut raisonner statistiquement comme il suit : si N_n est le nombre des atomes dans l'état d'énergie E_n, il y a une certaine probabilité $P_{nm}dt$ pour qu'un de ces atomes passe dans l'état d'énergie E_m pendant le temps dt et l'énergie rayonnée par seconde sera (en supposant N_n assez grand pour que sa décroissance soit négligeable) :

$$N_n\, P_{nm}\, (E_n - E_m) = N_n\, P_{nm}\, h\, \nu_{nm}\,. \tag{57}$$

Le problème est d'évaluer la probabilité P_{nm}. Pour cela, M. Bohr est parti d'une idée féconde : celle de correspondance. Il a admis que pour les états stationnaires très voisins qui correspondent aux grandes valeurs des nombres quantiques, on devait retrouver asymptotiquement les lois classiques et il a pu ainsi énoncer certaines règles pour la prévision des intensités et des polarisations. Le développement de la nouvelle Mécanique a permis de préciser les règles pressenties par M. Bohr.

Pour faire comprendre comment on peut parvenir aux nouvelles règles de prévision pour les intensités et les polarisations, nous commencerons par raisonner comme le fit d'abord M. Schrödinger. Considérons un système atomique à un électron dont la fonction d'onde soit :

$$\Psi = \sum_n c_n\, a_n\, (x, y, z)\, e^{\frac{2\pi i}{h} E_n t} \tag{58}$$

Nous avons vu qu'on pouvait lui associer la distribution d'électricité moyenne définie par la densité $\delta = -e\Psi\Psi^*$; le moment électrique correspondant est :

$$m_x = -e \int\int\int x \Psi\Psi^* d\tau$$

$$= C^{te} + \sum_{nm} (-e)\, c_n c_m^* \int\int\int x a_n a_m^* d\tau\, e^{\frac{2\pi i}{h}(E_n - E_m)t} \quad (59)$$

etc. Or l'intégrale qui figure au dernier membre de (59) est l'élément d'indices mn de la matrice engendrée par l'opérateur x. dans le système des a_i. Si on le désigne par X_{mn}, on peut écrire (59) sous la forme :

$$m_x = C^{te} + 2 \sum_{n,m}^{n \neq m} (-e) \cdot |c_n| \cdot |c_m| \cdot |X_{mn}| \cdot \cos 2\pi \left[\frac{E_n - E_m}{h} t + \varphi_m - \varphi_n \right] \cdot \quad (60)$$

Les fréquences qui apparaissent dans le développement (60) sont précisément les fréquences $\nu_{mn} = \dfrac{E_m - E_n}{h}$ de Bohr. Il est donc tout à fait naturel de penser qu'un ensemble d'atomes se trouvant dans l'état défini par la fonction d'onde (58) émettra la raie de fréquence ν_{mn} avec une intensité proportionnelle à $\nu_{mn}^4 | X_{mn} |^2$.

Mais, nous l'avons vu, le fait que les radiations sont émises par quanta liés aux transitions entre états stationnaires nous oblige à nous séparer de l'image classique d'une distribution électrique de densité $-e\Psi\Psi^*$ qui rayonnerait simultanément et continûment toutes les fréquences de Bohr. Nous sommes contraints, conformément à l'esprit même de la nouvelle Mécanique, d'adopter un énoncé purement statistique. Voici l'énoncé que nous devons adopter pour éviter toute contradiction : « Soit une assemblée d'atomes identiques pour lesquelles on a les fonctions propres a_n de l'opérateur Hamiltonien et par suite les composantes de matrice de la forme

$$X_{mn} = \int a_m^* x \cdot a_n d\tau ;$$

supposons que dans cette assemblée, il y ait N_n atomes dans l'état d'énergie E_n. La quantité d'énergie émise par seconde

par l'ensemble des atomes sous la forme d'un rayonnement de fréquence $\nu_{nm} = \frac{E_n - E_m}{h}$ ayant sa vibration électrique parallèle à l'axe des x est égale à $N_n \frac{64\,\pi^4}{3\,c^3} \nu_{nm}{}^4 \, |X_{mn}|^2$. »

En comparant avec (57), on voit que la probabilité $[P_{nm}]_x$ des transitions produisant le rayonnement en question est donnée par :

$$[P_{nm}]_x = \frac{64\,\pi^4}{3\,c^3} \frac{\nu_{nm}{}^3}{h} \left| X_{mn} \right|^2 \qquad (61)$$

Ceci est donc la probabilité de passage par unité de temps pour un atome de l'état E_n à l'état E_m avec émission d'une radiation dont la vibration soit parallèle à l'axe des x. Quand pour deux états E_n et E_m, les trois probabilités $[P_{nm}]_x$ $[P_{nm}]_y$ et $[P_{nm}]_z$ se trouvent nulles à la fois, l'émission du rayonnement de fréquence ν_{nm} n'a pas lieu réellement. On démontre qu'il en est ainsi pour toutes les transitions qui ne satisfont pas simultanément aux relations :

$$\delta l = \pm 1 \qquad\qquad \delta m = \begin{cases} 0 \\ \pm 1 \end{cases}. \qquad (62)$$

La Mécanique ondulatoire permet donc de justifier les règles de sélection pour les nombres quantiques l et m. Bien entendu, elle ne peut justifier la troisième règle de sélection, $\delta j = \begin{cases} 0 \\ \pm 1 \end{cases}$, que nous avons rencontrée puisqu'elle ignore le nombre j. Nous verrons que la théorie de Dirac permet de retrouver l'ensemble des trois règles de sélection.

DEUXIÈME PARTIE

LA THÉORIE DE L'ÉLECTRON MAGNÉTIQUE ET TOURNANT DE DIRAC
PRINCIPES GÉNÉRAUX

CHAPITRE IX

La tentative de Pauli

1. *Rotation de l'électron et polarisation de la lumière*

NOUS avons vu que la Mécanique ondulatoire sous sa forme primitive ne résout aucune des difficultés que l'on avait été amené à rattacher à l'existence d'un magnétisme propre à l'électron. Il est donc devenu évident que la Mécanique ondulatoire resterait incomplète tant qu'elle ne contiendrait pas un élément correspondant à ce magnétisme propre. Mais l'idée d'un électron magnétique et tournant ne peut pas se développer aussi facilement dans la nouvelle théorie que dans les anciennes. Restant dans le cadre des conceptions classiques, MM. Uhlenbeck et Goudsmit avaient imaginé l'électron comme une petite sphère d'électricité négative en rotation présentant un moment de rotation égal à $\frac{1}{2}.\frac{h}{2\pi}$ et un moment magnétique de même axe égal à un magnéton de Bohr $\frac{eh}{4\pi m_0 c}$. Mais en Mécanique ondulatoire, il paraît interdit de se faire une représentation aussi précise et l'on doit toujours introduire le langage des probabilités.

Pour se faire une idée de la façon dont on va avoir à poser le problème, il est bon de réfléchir à la manière dont on

doit définir avec les idées nouvelles la polarisation du quantum de lumière. Considérons un faisceau de lumière polarisé rectilignement qui se propage dans une direction *oz*.

Soient *ox* et *oy* deux axes perpendiculaires à *oz*. Le vecteur lumineux normal à *oz* est de la forme :

$$\overrightarrow{a_0} \sin 2\pi\nu\left(t - \frac{z}{c} + \varphi\right)$$

et il a des composantes suivant *ox* et *oy* dont les amplitudes sont données par les formules :

$$a_x = a_0 \cos\theta \qquad a_y = a_0 \sin\theta \tag{1}$$

si l'on nomme θ l'angle de ce vecteur avec *ox*. L'intensité lumineuse du faisceau est a_0^2. Si l'on place sur le trajet du faisceau un nicol qui laisse seulement passer la vibration parallèle à *ox*, le vecteur lumineux après le nicol se réduit à $a_x \sin 2\pi\nu\left(t - \frac{z}{c} + \varphi\right)$ et l'intensité est $a_0^2 \cos^2\theta$. Si ensuite on tourne le nicol de 90°, le vecteur lumineux dans la vibration transmise est parallèle à *oy* et égal à

$$a_y \sin 2\pi\nu\left(t - \frac{z}{c} + \varphi\right);$$

l'intensité transmise est $a_0^2 \sin^2\theta$. Comment doit-on interpréter tout ceci si l'on admet l'existence des photons? Si l'on attribue individuellement à chaque photon une polarisation, la description causale du phénomène devient impossible. En effet comme on est en quelque sorte contraint d'attribuer au photon incident la polarisation définie par le vecteur lumineux de l'onde incidente, on n'aperçoit pas comment il serait possible d'expliquer causalement pourquoi certains photons ne traversent pas le nicol alors que d'autres le traversent en prenant une polarisation dirigée suivant *ox* par exemple. En adoptant le point de vue probabiliste de la nouvelle Mécanique, on doit considérer que l'on ne peut pas attribuer au photon incident une polarisation bien déterminée mais que l'on doit seulement définir à l'aide de l'onde

lumineuse associée la probabilité pour que le photon en traversant le nicol se révèle comme ayant une polarisation parallèle à ox : cette probabilité est $\cos^2\theta$. On peut dire que le photon avant la traversée du nicol a une probabilité $\cos^2\theta$ de se révéler comme polarisé parallèlement à ox et une probabilité $\sin^2\theta$ de se révéler comme polarisé parallèlement à oy.

Revenons maintenant au spin de l'électron. L'hypothèse de l'électron magnétique et tournant conduit à associer à l'électron deux grandeurs dirigées proportionnelles l'une à l'autre, moment magnétique propre et moment de rotation propre. Tandis que pour le photon la polarisation est définie par une direction dépourvue de sens, le vecteur spin pour l'électron a une direction et un sens. Mais pas plus que nous ne pouvons attribuer à un photon une polarisation déterminée, nous ne devons pouvoir, dans une théorie de l'électron magnétique conforme aux principes généraux de la nouvelle Mécanique, attribuer un spin déterminé à un électron individuel. La seule chose dont nous pouvons parler, c'est de la probabilité pour qu'une expérience permettant de déterminer la direction du spin d'un électron donne tel ou tel résultat. C'est sur cette idée essentielle que M. Pauli a fondé une première tentative de théorie de l'électron magnétique dans le cadre de la Mécanique ondulatoire. Nous devons insister un peu sur cette tentative parce qu'elle a guidé M. Dirac dans le développement de sa théorie plus complète.

2. *Théorie de M. Pauli.* — M. Pauli (1) a fait une première tentative pour poser nettement le problème de l'électron magnétique dans le cadre des idées générales de la nouvelle Mécanique.

Si l'on considère un système d'axes rectangulaires, toute expérience permettant d'assigner une valeur à la composante du moment de rotation propre d'un électron le long

---

(1) *Zeitschrift für Physik.*, 43 (1927), p. 601.

de oz donnera pour résultat soit $+\frac{1}{2}\cdot\frac{h}{2\pi}$, soit $-\frac{1}{2}\cdot\frac{h}{2\pi}$ d'après la conception d'Uhlenbeck et Goudsmit. Dès lors il est naturel de supposer que l'on doit associer à l'électron non pas une fonction Ψ mais deux fonctions $\Psi_1(xyzt)$ *et* $\Psi_2(xyzt)$ de telle sorte que $|\Psi_1|^2\, dx\, dy\, dz$ mesure la probabilité pour qu'au temps t les coordonnées de l'électron soient comprises dans l'intervalle $dx\, dy\, dz$ et que la composante z de son moment de rotation soit $+\frac{1}{2}\cdot\frac{h}{2\pi}$ tandis que $|\Psi_2|^2\, dx\, dy\, dz$ mesure la probabilité pour qu'au temps t les coordonnées soient comprises dans l'intervalle $dx\, dy\, dz$, la composante z du moment de rotation étant $-\frac{1}{2}\cdot\frac{h}{2\pi}$. M. Pauli introduit ainsi l'idée que, pour tenir compte du magnétisme de l'électron, il faut augmenter le nombre des fonctions Ψ. Naturellement la condition de normalisation devient ici :

$$\int\int\int\Big[|\Psi_1|^2 + |\Psi_2|^2\Big]\, dx\, dy\, dz = 1\,. \qquad (2)$$

M. Pauli a cherché des équations de propagation déterminant les deux fonctions d'onde Ψ. Pour cela il part toujours de la fonction Hamiltonienne, mais en tenant compte du moment magnétique de l'électron.

Si on désigne par $s_x s_y s_z$ les composantes d'un vecteur unité $\vec{s}$ porté dans la direction du spin de l'électron, le moment magnétique est $\frac{eh}{4\pi m_0 c}\cdot\vec{s}$.

La fonction Hamiltonienne dans un champ constant sera alors de la forme $H(x, y, z, p_x, p_y, p_z, s_x, s_y, s_z)$ et les quantités $s_x s_y s_z$ y figureront linéairement comme il est aisé de le voir. Maintenant l'artifice qui permet à la Mécanique ondulatoire d'obtenir l'équation de propagation consiste à remplacer $p_x p_y p_z$ par les opérateurs $-\frac{h}{2\pi i}\cdot\frac{\partial}{\partial x}$, etc. et à écrire :

$$H\left(x, y, z, -\frac{h}{2\pi i}\frac{\partial}{\partial x}, -\frac{h}{2\pi i}\frac{\partial}{\partial y}, -\frac{h}{2\pi i}\frac{\partial}{\partial z}, s_x, s_y, s_z\right)\Psi = E\Psi \quad (3)$$

pour les ondes monochromatiques de fréquence $\frac{E}{h}$. Mais il nous faut deux équations pour les deux Ψ. Guidé par des considérations générales que je passe sous silence, Pauli imagine de remplacer les composantes $s_x\, s_y$ et s_z respectivement par les trois matrices hermitiques à 2 lignes et 2 colonnes :

$$\mathbf{s}_1 = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} \qquad \mathbf{s}_2 = \begin{vmatrix} 0 & -i \\ +i & 0 \end{vmatrix} \qquad \mathbf{s}_3 = \begin{vmatrix} +1 & 0 \\ 0 & -1 \end{vmatrix} \quad (4)$$

Il convient de plus de poser la définition suivante : A étant une matrice à 2 lignes et 2 colonnes, l'opération $A\Psi$ est défini par la relation :

$$A\Psi_i = \sum_{1}^{2}{}_k\, a_{ik}\, \Psi_k \qquad (i, k = 1, 2) \qquad (5)$$

généralisation naturelle de formules que nous avons déjà rencontrées à propos des matrices. L'application de la relation (5) aux matrices $\mathbf{s}_1\,\mathbf{s}_2\,\mathbf{s}_3$ donne en particulier :

$$\mathbf{s}_1\ \Psi_1 = \Psi_2\ ;\ \mathbf{s}_1\ \Psi_2 = \Psi_1\ ;\ \mathbf{s}_2\ \Psi_1 = -i\ \Psi_2\ ;\ \mathbf{s}_2\ \Psi_2 = i\ \Psi_1\ ;$$
$$\mathbf{s}_3\ \Psi_1 = \Psi_1\ ;\ \mathbf{s}_3\ \Psi_2 = -\Psi_2 \qquad (6)$$

Pauli remplace alors l'équation (3) par les deux équations:

$$\begin{aligned} H\left(x, y, z, -\frac{h}{2\pi i}\frac{\partial}{\partial x}, -\frac{h}{2\pi i}\frac{\partial}{\partial y}, -\frac{h}{2\pi i}\frac{\partial}{\partial z}, \mathbf{s}_1, \mathbf{s}_2\mathbf{s}_3\right)\Psi_1 = E\Psi_1 \\ H\left(x, y, z, -\frac{h}{2\pi i}\frac{\partial}{\partial x}, -\frac{h}{2\pi i}\frac{\partial}{\partial y}, -\frac{h}{2\pi i}\frac{\partial}{\partial z}, \mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3\right)\Psi_2 = E\Psi_2 \end{aligned} \quad (7)$$

Ce sont deux équations simultanées auxquelles doivent satisfaire les fonctions Ψ_1 et Ψ_2.

Mais la théorie développee plus haut fait jouer un rôle particulier à l'axe *oz* puisque les fonctions d'onde Ψ_1 et Ψ_2 donnent les probabilités des deux valeurs possibles pour la composante du spin *parallèle à* oz. Si l'on veut obtenir les probabilités pour que la composante du moment de rotation parallèle à un autre axe quelconque OZ se révèle égale à $\frac{1}{2} \cdot \frac{h}{2\pi}$ ou à $-\frac{1}{2} \cdot \frac{h}{2\pi}$, il faut prendre un autre

système d'axes rectangulaires OXYZ dans lequel OZ soit le troisième axe et calculer dans ce système les deux fonctions d'onde, disons $\Phi_1(X, Y, Z)$ et $\Phi_2(X, Y, Z)$, dont les carrés des modules donneront les probabilités cherchées et naturellement, comme rien ne distingue physiquement le système d'axes OXYZ du système $oxyz$, les fonctions Φ devront obéir aux équations suivantes que l'on obtient par transposition immédiate de (7):

$$\begin{aligned} &H\left(X, Y, Z, -\frac{h}{2\pi i}\frac{\partial}{\partial X}, -\frac{h}{2\pi i}\frac{\partial}{\partial Y}, -\frac{h}{2\pi i}\frac{\partial}{\partial Z}, \mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3\right)\Phi_1 = E\Phi_1 \\ &H\left(X, Y, Z, -\frac{h}{2\pi i}\frac{\partial}{\partial X}, -\frac{h}{2\pi i}\frac{\partial}{\partial Y}, -\frac{h}{2\pi i}\frac{\partial}{\partial Z}, \mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3\right)\Phi_2 = E\Phi_2 . \end{aligned} \tag{8}$$

La position des axes OXYZ par rapport aux axes $oxyz$ peut se définir de la façon suivante. Remplaçons x, y, z, X, Y, Z par $x_1, x_2, x_3, X_1, X_2, X_3$, nous avons des formules de transformation de la forme :

$$X_i = \sum_j o_{ij} x_j \tag{9}$$

les o_{ij} étant les éléments d'une matrice réelle et orthogonale $\left(\sum_i o_{ij}\, o_{ik} = \delta_{jk}\right)$. En somme la matrice O définit le passage du premier système d'axes au second. Cette matrice étant connue, comment peuvent s'exprimer les fonctions Φ à l'aide des fonctions Ψ? Telle est la question que Pauli s'est posée et qu'il a résolue. Si l'on part des équations (7) et si l'on remplace xyz en fonction de XYZ par les formules inverses de (9), on obtient, comme nous le vérifierons plus loin dans un cas particulier, les équations :

$$H\left(X, Y, Z, -\frac{h}{2\pi i}\frac{\partial}{\partial X}, -\frac{h}{2\pi i}\frac{\partial}{\partial Y}, -\frac{h}{2\pi i}\frac{\partial}{\partial Z}, S_1, S_2, S_3\right)\Psi_1(X, Y, Z) = E\,\Psi_1(X, Y, Z) \tag{10}$$

$$H\left(X, Y, Z, -\frac{h}{2\pi i}\frac{\partial}{\partial X}, -\frac{h}{2\pi i}\frac{\partial}{\partial Y}, -\frac{h}{2\pi i}\frac{\partial}{\partial Z}, S_1, S_2, S_3\right)\Psi_2(X, Y, Z) = E\,\Psi_2(X, Y, Z)$$

L'opérateur H des équations (10) est défini de la manière suivante : on écrit la fonction Hamiltonienne

$$H(X, Y, Z, p_X, p_Y, p_Z, s_X, s_Y, s_Z)$$

classique dans le système XYZ, fonction dans laquelle figurent les composantes s_X s_Y et s_Z du vecteur unité porté suivant le moment magnétique propre de l'électron ; puis l'on remplace dans cette expression d'une part p_X, p_Y, p_Z par $-\frac{h}{2\pi i}\frac{\partial}{\partial X}$ etc. et d'autre part s_X s_Y s_Z par les matrices hermitiques S_1 S_2 et S_3 définies par les relations :

$$S_i = \sum_j \alpha_{ij} \mathbf{s}_j \tag{11}$$

à partir des $\mathbf{s}_i$ données par (4). Dans l'opérateur H ainsi formé, les S_i figurent toujours linéairement. Enfin dans les formules (10) les fonctions $\Psi_1(X, Y, Z)$ et $\Psi_2(X, Y, Z)$ sont obtenues à partir des $\Psi_1(x, y, z)$ et $\Psi_2(x, y, z)$ en remplaçant xyz par leurs expressions en X, Y, Z.

M. Pauli a alors démontré, et nous verrons une démonstration analogue dans la théorie de Dirac ([1]), qu'il existe toujours une matrice unitaire Λ à deux lignes et deux colonnes telle que l'on ait :

$$S_i = \Lambda^{-1}\mathbf{s}_i\Lambda \qquad (\Lambda^{-1} = \Lambda^{+}) \quad i = 1, 2, 3 \tag{12}$$

En d'autres termes, il existe une transformation canonique faisant passer de chacun des $\mathbf{s}_i$ au S_i correspondant. A la matrice Λ correspond une opération définie par la formule (5) et en appliquant cette opération aux deux membres des équations (10), on obtient :

([1]) Cf. chapitre XI, paragraphe 2.

$$\Lambda H\left(X,Y,Z,-\frac{h}{2\pi i}\frac{\partial}{\partial X},-\frac{h}{2\pi i}\frac{\partial}{\partial Y},-\frac{h}{2\pi i}\frac{\partial}{\partial Z},S_1,S_2,S_3\right)\Psi_1=E\Lambda\Psi_1$$

$$\Lambda H\left(X,Y,Z,-\frac{h}{2\pi i}\frac{\partial}{\partial X},-\frac{h}{2\pi i}\frac{\partial}{\partial Y},-\frac{h}{2\pi i}\frac{\partial}{\partial Z},S_1,S_2,S_3\right)\Psi_2=E\Lambda\Psi_2 \qquad (13)$$

les Ψ étant exprimées en XYZ. Comme les S_i figurent linéairement dans H, Λ H contient linéairement les termes Λ S_i égaux d'après (12) à $\mathbf{s}_i\Lambda$; comme de plus Λ permute avec X, $\frac{\partial}{\partial X}$ etc., on a :

$$\Lambda H(\ldots)\Psi_1=H\left(X,Y,Z,-\frac{h}{2\pi i}\frac{\partial}{\partial X},-\frac{h}{2\pi i}\frac{\partial}{\partial Y},-\frac{h}{2\pi i}\frac{\partial}{\partial Z},\mathbf{s}_1,\mathbf{s}_2,\mathbf{s}_3\right)\Lambda\Psi_1=E\Lambda\Psi_1 \qquad (14)$$

et une équation analogue en Ψ_2. En comparant avec (8), on voit que l'on a :

$$\Phi_1(X,Y,Z)=\Lambda\Psi_1(X,Y,Z);\ \Phi_2(X,Y,Z)=\Lambda\Psi_2(X,Y,Z). \qquad (15)$$

Ces formules (15) nous apprennent comment les fonctions d'onde Φ se déduisent des fonctions d'onde Ψ car, dans chaque cas particulier, la matrice O étant connue, on peut trouver la matrice Λ. Nous allons en donner un exemple concret au paragraphe suivant.

Bien entendu, Λ étant unitaire, la somme $|\Phi_1|^2+|\Phi_2|^2$ est égale à $|\Psi_1|^2+|\Psi_2|^2$, ce qui exprime le fait que la probabilité totale des deux hypothèses possibles sur le sens du moment de rotation est toujours l'unité.

3. *Exemple d'application de la théorie précédente.* — Pour illustrer la théorie du précédent paragraphe, il est utile d'étudier un exemple très simple développé par Pauli lui-même :

celui d'un électron au repos dans un champ magnétique $\vec{H}$. Prenons d'abord un système d'axes $oxyz$ tel que la direction positive de oz coïncide avec la direction de $\vec{H}$. On a alors $H_x = H_y = 0$, $H_z = H$. En introduisant le magnéton de Bohr $\mu_0 = \frac{eh}{4\pi m_0 c}$, la fonction Hamiltonienne (¹) se réduit à $\mu_0 s_z H$ et l'on a pour les équations (7):

$$\mu_0 \mathbf{s}_3 H\Psi_1 = E\Psi_1 \qquad \mu_0 \mathbf{s}_3 H\Psi_2 = E\Psi_2 \tag{16}$$

ou, à cause de (6):

$$\mu_0 H\Psi_1 = E\Psi_1 \qquad -\mu_0 H\Psi_2 = E\Psi_2 \tag{16bis}$$

Ce système n'a de solution que pour $E = \pm \mu_0 H$. Comme $|\Psi_1|^2 + |\Psi_2|^2 = 1$, on a :

$$\begin{aligned} &\text{pour } E = \mu_0 H : && \Psi_1 = e^{i\gamma} && \Psi_2 = 0 \\ &\text{pour } E = -\mu_0 H : && \Psi_1 = 0 && \Psi_2 = e^{i\delta} \end{aligned} \tag{17}$$

Donc pour $E = \mu_0 H$, l'axe magnétique de l'électron pointe certainement vers les z positifs et pour $E = -\mu_0 H$, il pointe certainement vers les z négatifs. Ce sont bien les résultats auxquels on pouvait s'attendre.

Prenons maintenant un second système d'axes OXYZ tel que l'axe OZ fasse un angle θ avec le champ $\vec{H}$. Alors la fonction Hamiltonienne classique est

$$\mu_0 (\vec{s} \cdot \vec{H}) = \mu_0 (s_X H_X + s_Y H_Y + s_Z H_Z)$$

et les fonctions d'onde Φ_1 et Φ_2 sont solutions, d'après (8), des équations :

$$\begin{aligned} \mu_0 [H_X \mathbf{s}_1 + H_Y \mathbf{s}_2 + H_Z \mathbf{s}_3] \Phi_1 &= E\Phi_1 , \\ \mu_0 [H_X \mathbf{s}_1 + H_Y \mathbf{s}_2 + H_Z \mathbf{s}_3] \Phi &= E\Phi_2 \end{aligned} \tag{18}$$

(¹) Dans ce paragraphe la lettre H ne désigne plus l'Hamiltonien, mais la grandeur du champ magnétique.

ce qui donne en explicitant :

$$\begin{aligned} \mu_0 \left[(H_x - iH_y)\, \Phi_2 + H_z\, \Phi_1\right] &= E\Phi_1\ ; \\ \mu_0 \left[(H_x + iH_y)\, \Phi_1 - H_z\, \Phi_2\right] &= E\Phi_2 \end{aligned} \tag{19}$$

Pour que ces deux équations homogènes en Φ_i soient compatibles, il faut que :

$$\begin{vmatrix} \mu_0\, H_z - E & \mu_0\, (H_x - i\, H_y) \\ \mu_0\, (H_x + i\, H_y) & -\mu_0\, H_z - E \end{vmatrix} = 0 \tag{20}$$

c'est-à-dire

$$E = \pm\, \mu_0 \sqrt{H_x^2 + H_y^2 + H_z^2} = \pm\, \mu_0\, H\,, \tag{21}$$

Deux cas sont donc ici encore à distinguer :

Premier cas : $E = \mu_0\, H$.

Alors on a :

$$\frac{\Phi_2}{\Phi_1} = \frac{H - H_z}{H_x - i\, H_y} = \frac{H - H_z}{H_x^2 + H_y^2}\,(H_x - i\, H_y) = \frac{H - H_z}{H_p^2} \cdot H_p e^{i\alpha} \tag{22}$$

en posant $\alpha = \text{arc tg}\, \frac{H_y}{H_x}$ et en désignant par H_p la composante de $\vec{H}$ normale à OZ. On a donc aussi :

$$\frac{\Phi_2}{\Phi_1} = \frac{H\,(1 - \cos\theta)}{H \sin\theta}\, e^{i\alpha} = \text{tg}\, \frac{\theta}{2} \cdot e^{i\alpha} \tag{23}$$

Comme $|\Phi_1|^2 + |\Phi_2|^2 = 1$, on doit avoir :

$$|\Phi_1|^2 = \frac{1}{1 + \text{tg}^2\, \frac{\theta}{2}} = \cos^2 \frac{\theta}{2}\ ;\quad |\Phi_2|^2 = \sin^2 \frac{\theta}{2}. \tag{24}$$

Comme les arguments sont arbitraires, nous pouvons poser :

$$\Phi_1 = \cos \frac{\theta}{2}\, e^{-i \frac{\alpha - \frac{\pi}{2}}{2}}\ ;\quad \Phi_2 = i \sin \frac{\theta}{2}\, e^{i \frac{\alpha - \frac{\pi}{2}}{2}} \tag{25}$$

Deuxième cas : $E = -\mu_0 H$.

On a alors :

$$\frac{\Phi_2}{\Phi_1} = -\frac{H + H_x}{H_x - i H_y} =$$

$$-\frac{H + H_x}{H_p} \cdot e^{i\alpha} = -\frac{1+\cos\theta}{\sin\theta} e^{i\alpha} = -\operatorname{ctg}\frac{\theta}{2} \cdot e^{i\alpha} \qquad (26)$$

et par suite :

$$|\Phi_1|^2 = \frac{1}{1 + \operatorname{ctg}^2\frac{\theta}{2}} = \sin^2\frac{\theta}{2} \; ; \quad |\Phi_2|^2 = \cos^2\frac{\theta}{2} \qquad (27)$$

ce qui permet de poser :

$$\Phi_1 = i \sin\frac{\theta}{2} \cdot e^{-i\frac{\alpha - \frac{\pi}{2}}{2}} \; ; \quad \Phi_2 = \cos\frac{\theta}{2} \cdot e^{i\frac{\alpha - \frac{\pi}{2}}{2}} \qquad (28)$$

Les résultats ainsi obtenus signifient, pour prendre un exemple, que si une première expérience nous a montré que l'électron pointait dans la direction du champ avec l'énergie $\mu_0 H$, la probabilité pour qu'une seconde expérience conduise à nous le montrer pointant dans la direction OZ est $\cos^2\frac{\theta}{2}$.

Dans l'exemple simple traité ici, il est facile de calculer les matrices unitaires Λ de Pauli. Comparons les fonctions Ψ de (17) (en y prenant nuls pour simplifier les arguments arbitraires γ et δ) avec les fonctions Φ de (25) et de (28). Nous voyons que la matrice unitaire Λ est égale à :

$$\Lambda = \begin{vmatrix} \cos\frac{\theta}{2} \cdot e^{-i\frac{\alpha - \frac{\pi}{2}}{2}} & i \sin\frac{\theta}{2} \cdot e^{-i\frac{\alpha - \frac{\pi}{2}}{2}} \\ i \sin\frac{\theta}{2} \, e^{i\frac{\alpha - \frac{\pi}{2}}{2}} & \cos\frac{\theta}{2} \, e^{i\frac{\alpha - \frac{\pi}{2}}{2}} \end{vmatrix} \qquad (29)$$

Les éléments Λ_{ij}, qui sont des cas particuliers des paramètres de Cayley-Klein, satisfont la relation :

$$\Lambda_{11}\Lambda_{22} - \Lambda_{12}\Lambda_{21} = 1 \tag{30}$$

On trouve aisément

$$\Lambda^{+} = \begin{vmatrix} \Lambda_{22} & -\Lambda_{12} \\ -\Lambda_{21} & \Lambda_{11} \end{vmatrix} \tag{31}$$

d'où l'on tire :

$$\Lambda\Lambda^{+} = \begin{vmatrix} \Lambda_{22}\Lambda_{11} - \Lambda_{12}\Lambda_{21} & -\Lambda_{11}\Lambda_{12} + \Lambda_{12}\Lambda_{11} \\ \Lambda_{21}\Lambda_{22} - \Lambda_{22}\Lambda_{21} & \Lambda_{22}\Lambda_{11} - \Lambda_{12}\Lambda_{21} \end{vmatrix} = 1 \tag{32}$$

Λ est donc bien unitaire. La forme de Λ pouvait d'ailleurs être prévue par les raisonnements de Pauli.

4. ***Insuffisance de la théorie de Pauli.*** — Nous avons seulement donné une esquisse de la théorie de Pauli. Cette théorie n'est pas en effet entièrement satisfaisante. D'abord elle n'est pas en accord avec le principe de Relativité : elle n'envisage que les changements de coordonnées spatiales et non pas les transformations de coordonnées d'espace-temps au sens relativiste. De plus, elle ne conduit pas à une prévision tout à fait correcte du spectre de l'hydrogène.

Pour ces raisons nous ne l'étudierons pas davantage, mais nous devons remarquer qu'elle a introduit les idées essentielles suivantes : 1° le magnétisme de l'électron correspond à l'existence de plusieurs fonctions Ψ ; 2° les fonctions d'onde Ψ doivent permettre de définir la probabilité des orientations possibles du spin dans une certaine direction ; 3° en changeant les axes de coordonnées, on peut conserver des équations de propagation de même forme, mais les fonctions d'onde se transforment alors d'une certaine manière définie par une matrice Λ.

M. Darwin [1] a fait une autre tentative pour introduire le magnétisme de l'électron en Mécanique ondulatoire d'une

---

[1] *Proceeding Royal Society*, A, 116 (1927), p. 227.

façon conforme au principe de Relativité : il a cherché à définir 4 fonctions Ψ qui seraient les composantes d'un vecteur d'espace-temps. La tentative n'a pas pleinement réussi et il s'est ensuite lui-même rallié à la théorie de Dirac où figurent aussi 4 fonctions Ψ mais qui ne sont pas les composantes d'un vecteur d'Univers.

CHAPITRE X

La théorie de Dirac

1. *Rappel des résultats antérieurs*

POUR introduire dans les équations de la Mécanique ondulatoire l'idée du magnétisme et de la rotation propres de l'électron, M. Dirac est parti de considérations plus générales que celles de Pauli. Il a admis que les équations de la Mécanique ondulatoire devaient être mises en accord avec le principe de relativité restreinte, mais il a critiqué la façon dont on avait cherché à effectuer cette mise en accord. Pour comprendre ses critiques, rappelons d'abord rapidement comment nous avons écrit les équations de la Mécanique ondulatoire à un Ψ.

L'équation générale de propagation pour un corpuscule dans la Mécanique ondulatoire non relativiste est :

$$H\left(x, y, z, -\frac{h}{2\pi i}\frac{\partial}{\partial x}, -\frac{h}{2\pi i}\frac{\partial}{\partial y}, -\frac{h}{2\pi i}\frac{\partial}{\partial z}, t\right)\Psi = \frac{h}{2\pi i}\frac{\partial \Psi}{\partial t} \qquad (1)$$

H étant l'opérateur Hamiltonien. Dans le cas d'un corpuscule de masse m se déplaçant dans un champ caractérisé par un potentiel $U(x, y, z, t)$, on a :

$$-\frac{h^2}{8\pi^2 m}\Delta\Psi + U(x, y, z, t)\Psi = \frac{h}{2\pi i}\frac{\partial \Psi}{\partial t} \qquad (2)$$

Si le corpuscule porte une charge ε et se meut dans un champ électrostatique dérivant du potentiel électrique $V(x, y, z, t)$, on a $U = \varepsilon V$ et (2) s'écrit :

$$-\frac{h^2}{8\pi^2 m} \Delta \Psi + \varepsilon V(x\,y\,z\,t)\,\Psi = \frac{h}{2\pi i}\frac{\partial \Psi}{\partial t} \qquad (3)$$

En particulier si le corpuscule chargé est un électron, on doit poser dans (3) : $\varepsilon = -e$

Nous avons vu que dans cette théorie non relativiste la quantité nécessairement positive $\Psi\Psi^*$ représentait la densité de probabilité de présence et que la probabilité totale ainsi définie se conservait (égale à l'unité) au cours du temps. La quantité $\varepsilon\Psi\Psi^*$ ($-e\Psi\Psi^*$ pour l'électron) est la densité électrique moyenne qui sert au calcul du rayonnement émis.

Lorsqu'on a voulu construire une Mécanique ondulatoire en accord avec le principe de Relativité restreinte, on a obtenu tout de suite par une induction très naturelle une nouvelle équation de propagation pour remplacer (1). Nous avons vu que cette nouvelle équation s'écrit pour un corpuscule de masse m et de charge ε se déplaçant dans un champ électromagnétique défini par un potentiel scalaire $V(xyzt)$ et un potentiel vecteur $\vec{A}(xyzt)$:

$$\frac{1}{c^2}\left(\frac{h}{2\pi i}\frac{\partial}{\partial t} - \varepsilon V\right)^2 \Psi - \sum_{xyz}\left(\frac{h}{2\pi i}\frac{\partial}{\partial x} + \frac{\varepsilon}{c}A_x\right)^2 \Psi = m_0^2 c^2 \Psi \qquad (4)$$

Pour un électron, l'équation (4) prend la forme :

$$\frac{1}{c^2}\left(\frac{h}{2\pi i}\frac{\partial}{\partial t} + eV\right)^2 \Psi - \sum_{xyz}\left(\frac{h}{2\pi i}\frac{\partial}{\partial x} - \frac{e}{c}A_x\right)^2 \Psi = m_0^2 c^2 \Psi. \qquad (5)$$

Introduisons les notations :

$$P_1 = -\frac{h}{2\pi i}\frac{\partial}{\partial x} + \frac{eA_x}{c}\,;\quad P_2 = -\frac{h}{2\pi i}\frac{\partial}{\partial y} + \frac{eA_y}{c}\,;$$
$$P_3 = -\frac{h}{2\pi i}\frac{\partial}{\partial z} + \frac{eA_z}{c}\,;\quad P_4 = \frac{h}{2\pi i}\frac{1}{c}\frac{\partial}{\partial t} + \frac{eV}{c} \qquad (6)$$

Elles nous permettent de remplacer l'équation (5) par :

$$\left(P_4^2 - \sum_1^3 P_i^2 - m_0^2 c^2\right)\Psi = 0 \tag{7}$$

Les équations de propagation relativistes (4)-(7) ne nous permettent plus, nous l'avons vu, de prendre pour densité de la probabilité de localisation la quantité $\Psi\Psi^*$ parce que l'intégrale de cette quantité dans tout l'espace n'est plus nécessairement constante au cours du temps. Nous avons dû adopter pour densité de probabilité de présence une expression plus compliquée qui dans le cas de l'électron ($\varepsilon = -e$) s'écrit :

$$\rho = \frac{-h}{4\pi m_0 c^2}\left(\Psi \cdot \frac{\partial \Psi^*}{\partial t} - \Psi^* \frac{\partial \Psi}{\partial t}\right) + \frac{e}{m_0 c^2} V \Psi \Psi^* \tag{8}$$

expression qui se réduit à $\Psi\Psi^*$ quand l'approximation Newtonienne est suffisante. La densité électrique moyenne est ici $\rho\varepsilon = -\rho e$ avec la valeur (8) de ρ.

2. ***Critiques adressées par M. Dirac aux équations 4-8.*** — M. Dirac a adressé de graves critiques à la tentative de Mécanique ondulatoire relativiste que nous venons de rappeler. Il s'est attaqué en particulier à l'expression (8) de la densité de probabilité ; comme cette expression est imposée par la forme même de l'équation de propagation (7), ceci revient à attaquer cette équation elle-même.

Résumons les arguments de Dirac. Une première critique que l'on peut adresser à l'expression (8) de ρ, c'est qu'elle n'est pas nécessairement positive alors qu'une valeur négative de ρ ne peut avoir évidemment aucun sens. Une autre critique déjà signalée à la fin du chapitre VII est la suivante : les principes généraux de la nouvelle Mécanique exigent, quelle que soit la forme de l'équation de propagation adoptée, que la probabilité de trouver, pour les coordonnées du corpuscule, des valeurs comprises dans les intervalles $x \rightarrow x + dx$, $y \rightarrow y + dy$, $z \rightarrow z + dz$ soit $\Psi\Psi^* \, dx \, dy \, dz$ et cette exigence ne cadre pas avec la formule (8).

En présence de ces difficultés, M. Dirac maintient que la densité de probabilité de localisation doit avoir nécessairement la forme toujours positive $\Psi\Psi^*$ ou, si l'on admet avec Pauli l'existence de plusieurs fonctions d'onde Ψ_i, la forme également toujours positive $\sum_i \Psi_i \Psi_i^*$. Mais, en adoptant ce postulat, il est amené à une conclusion d'une importance capitale : l'équation ou les équations de propagation de la véritable Mécanique ondulatoire relativiste doivent être du premier ordre par rapport aux quatre variables *xyzt*. Exposons son raisonnement.

L'équation non relativiste (1) est du premier ordre par rapport au temps et l'expression correspondante de la densité de probabilité est $\Psi\Psi^*$. Si on se donne la forme initiale $\Psi(xyzo)$ de la fonction d'onde, celle-ci se trouve entièrement déterminée par l'équation de propagation et par suite on peut dire qu'en se donnant le $\Psi\Psi^*$ initial l'évolution ultérieure de la densité est déterminée : on comprend donc que la conservation de la probabilité (et de l'électricité) soit alors une conséquence nécessaire de l'équation de propagation elle-même, comme nous l'avons prouvé. Dans le cas de l'équation relativiste (4), il n'en est plus de même ; cette équation étant du second ordre en t, il faut en effet se donner les valeurs de Ψ et de $\frac{\partial \Psi}{\partial t}$ à l'instant initial pour que la fonction d'onde soit déterminée.

C'est alors l'expression (8) qu'il faut adopter pour la densité de probabilité si l'on veut que la conservation de la probabilité résulte de l'équation de propagation. Mais adoptons maintenant avec Dirac le postulat que la densité de probabilité est nécessairement de la forme $\Psi\Psi^*$ ou $\sum_i \Psi_i \Psi_i^*$. Nous allons voir que l'équation (ou les équations de propagation) doit alors être du premier ordre en t et que par suite l'équation (4) ne peut être exacte.

En effet, si l'équation de propagation est du second ordre

en t, sa solution n'est déterminée que si l'on se donne les valeurs initiales de Ψ et de $\frac{\partial \Psi}{\partial t}$.

Supposons qu'on se donne seulement le Ψ initial mais pas le $\frac{\partial \Psi}{\partial t}$. Dans l'hypothèse de Dirac la densité initiale serait ainsi connue puisqu'elle ne dépend que de Ψ, mais l'évolution ultérieure de cette densité ne serait pas déterminée puisque celle de la fonction Ψ ne le serait pas ; en se donnant une forme arbitraire du $\frac{\partial \Psi}{\partial t}$ initial, on pourrait, en partant d'une forme initiale connue de la densité, avoir n'importe quelle évolution ultérieure pour cette densité et il ne pourrait y avoir conservation automatique de la probabilité. Dirac conclut donc que l'équation de propagation est du premier ordre en t et, comme le principe de Relativité fait toujours jouer un rôle symétrique aux coordonnées d'espace et de temps, elle doit être du premier ordre par rapport aux quatre variables $xyzt$. Le même raisonnement et la même conclusion sont valables s'il y a plusieurs fonctions Ψ et plusieurs équations de propagation simultanées.

M. Dirac devait donc chercher une ou plusieurs équations du premier ordre par rapport aux variables $xyzt$ pour remplacer l'équation (4). Nous allons voir comment il y a réussi.

3. ***Les équations de Dirac en l'absence de champ.*** — Pour établir les équations nouvelles, Dirac a commencé par le cas du mouvement libre de l'électron en l'absence de tout champ électromagnétique.

On a alors $V = \vec{A} = 0$ et les opérateurs P_i (formules 6) deviennent :

$$p_1 = -\frac{h}{2\pi i}\frac{\partial}{\partial x}; \; p_2 = -\frac{h}{2\pi i}\frac{\partial}{\partial y}; \; p_3 = -\frac{h}{2\pi i}\frac{\partial}{\partial z}; \; p_4 = \frac{h}{2\pi i}\frac{1}{c}\frac{\partial}{\partial t} \tag{9}$$

Dans ce cas simple, on peut admettre avec certitude que la fonction Ψ satisfait à l'équation :

$$\Delta \Psi - \frac{1}{c^2} \frac{\partial^2 \Psi}{\partial t^2} = \frac{4\pi^2}{h^2} m_0^2 c^2 \Psi \quad \text{ou} \left(p_4^2 - \sum_{i}^{3}{}_{1} p_i^2 - m_0^2 c^2 \right) \Psi = 0 \tag{10}$$

car en ce cas il résulte des considérations fondamentales dont est sortie la Mécanique ondulatoire que l'onde plane monochromatique de fréquence $\frac{W}{h} = \frac{m_0 c^2}{h\sqrt{1-\beta^2}}$ et de longueur d'onde $\lambda = \frac{h}{p} = \frac{h\sqrt{1-\beta^2}}{m_0 v}$ doit être solution et l'on vérifie sans peine qu'il en est ainsi avec l'équation (10).

Mais l'équation (10) est du second ordre et nous voulons des équations du premier ordre. Pour y parvenir, Dirac suppose qu'il existe plusieurs fonctions d'ondes Ψ_1, Ψ_2 ... Ψ_N et que chacune d'elles satisfait à l'équation :

$$\left(p_4^2 - \sum_{1}^{3}{}_{i} p_i^2 - m_0^2 c^2 \right) \Psi_k = 0 \quad (k=1, \ldots N) \tag{11}$$

mais ces équations du second ordre doivent être des *conséquences* des véritables équations de propagation qui sont du premier ordre.

Ces véritables équations de propagation du premier ordre, M. Dirac les écrit sous la forme symbolique :

$$(p_4 + \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \alpha_4 m_0 c) \Psi = 0 \tag{12}$$

L'équation symbolique (12) signifie que l'on a pour chaque Ψ_k :

$$\left(p_4 + \sum_{1}^{3}{}_{i} \alpha_i p_i + \alpha_4 m_0 c \right) \Psi_k = 0 . \tag{13}$$

Les α_i sont des matrices à N rangs et N colonnes et l'on définit l'opération $\alpha_i \Psi_k$ comme nous l'avons déjà fait dans l'étude de la théorie de Pauli par la formule :

$$\alpha_i \Psi_k = \sum_{1}^{N}{}_{l} \alpha_{i,kl} \Psi_l \tag{14}$$

où $\alpha_{i,kl}$ désigne l'élément d'indices kl de la matrice α_i.

Mais les équations (13) doivent avoir pour conséquence les équations (11) et cela impose certaines conditions aux matrices α_i. En effet appliquons à l'équation (13) l'opérateur $p_4 - \sum_1^3 {}_i \, \alpha_i \, p_i - \alpha_4 \, m_0 c$; nous trouvons :

$$\left[p_4^2 - \left(\sum_1^3 {}_i \, \alpha_i \, p_i + \alpha_4 \, m_0 c\right)\right] \Psi = 0 \qquad (15)$$

et l'équation (15) ne coïncide avec (11) que si l'on pose :

$$\alpha_i^2 = 1 \qquad \alpha_i \, \alpha_j + \alpha_j \, \alpha_i = 0. \qquad (16)$$

Nous devons donc imposer aux matrices α_i les conditions (16). De plus nous leur imposerons aussi d'être hermitiques comme toutes les matrices qui interviennent dans la nouvelle Mécanique.

M. Dirac a cherché à prendre le nombre N des fonctions d'onde le plus petit possible. Pour $N < 4$, on ne peut pas trouver 4 matrices hermitiques satisfaisant aux conditions (16). Pour $N = 4$, au contraire, il est possible d'en trouver et il sera commode d'adopter les matrices α_i suivantes :

$$\alpha_1 = \begin{vmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{vmatrix} \quad \alpha_2 = \begin{vmatrix} 0 & 0 & 0 & +i \\ 0 & 0 & -i & 0 \\ 0 & +i & 0 & 0 \\ -i & 0 & 0 & 0 \end{vmatrix} \quad \alpha_3 = \begin{vmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{vmatrix} \quad \alpha_4 = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{vmatrix} \qquad (17)$$

Elles sont visiblement hermitiques et de plus elles ont pour carrés la matrice 1 et anticommutent entre elles comme le veulent les conditions (16).

Bref, nous allons admettre avec Dirac l'existence de 4 fonctions Ψ_i qui satisfont aux 4 équations simultanées du premier ordre (13). Explicitement ces 4 équations s'écrivent avec le choix (17) des matrices α_i :

$$\begin{aligned} (p_4 + m_0 c) \Psi_1 + (p_1 + i p_2) \Psi_4 + p_3 \Psi_3 &= 0 \\ (p_4 + m_0 c) \Psi_2 + (p_1 - i p_2) \Psi_3 - p_3 \Psi_4 &= 0 \\ (p_4 - m_0 c) \Psi_3 + (p_1 + i p_2) \Psi_2 + p_3 \Psi_1 &= 0 \\ (p_4 - m_0 c) \Psi_4 + (p_1 - i p_2) \Psi_1 - p_3 \Psi_2 &= 0. \end{aligned} \qquad (18)$$

En partant du système (18), on peut voir d'une façon détaillée comment ce système a pour conséquence le système du second ordre (11) (avec $N=4$).

Appliquons par exemple à la première équation (18) l'opérateur $p_4 - m_0 c$ et à la quatrième équation (18) l'opérateur $p_1 + ip_2$ qui permute avec le précédent ; nous obtenons :

$$\begin{aligned} &(p_4^2 - m_0^2c^2)\Psi_1 + (p_4 - m_0c)(p_1 + ip_2)\Psi_4 + (p_4 - m_0c)p_3\Psi_3 = 0 \\ &(p_1 + ip_2)(p_4 - m_0c)\Psi_4 + (p_1^2 + p_2^2)\Psi_1 - (p_1 + ip_2)p_3\Psi_2 = 0 \end{aligned} \tag{19}$$

d'où résulte par élimination de Ψ_4 :

$$\begin{aligned} (p_4^2 - m_0^2c^2)\Psi_1 - (p_1^2 + p_2^2)\Psi_1 \\ + (p_1 + ip_2)p_3\Psi_2 + (p_4 - m_0c)p_3\Psi_3 = 0. \end{aligned} \tag{20}$$

Mais en appliquant l'opérateur p_3 à la troisième équation (18), on trouve :

$$p_3(p_4 - m_0c)\Psi_3 = -p_3(p_1 + ip_2)\Psi_2 - p_3^2\Psi_1 \tag{21}$$

et comme les p_i permutent entre eux, on obtient en substituant dans (20) l'équation :

$$\left(p_4^2 - \sum_1^3{}_i\, p_i^2 - m_0^2c^2\right)\Psi_1 = 0 \tag{22}$$

et l'on trouverait de même les équations du second ordre relatives aux trois autres Ψ. On a ainsi vérifié un résultat certain d'avance puisque les matrices (17) satisfont les conditions (16).

Les équations (18) de Dirac présentent un aspect très dissymétrique : l'axe des z et l'opérateur p_3 correspondant y jouent visiblement un rôle particulier.

Les fonctions Ψ_i solutions de ces équations sont donc intimement liées au choix des axes comme dans la théorie de Pauli ; elles doivent servir à calculer des probabilités pour lesquels l'axe des z joue un rôle particulier. Si l'on prend d'autres axes, on pourra écrire des équations de propagation ayant toujours la forme (18), mais elles auront pour solutions d'autres fonctions Ψ_i reliées aux précédentes par des

transformations analogues à celles que nous avons rencontrées dans la théorie de Pauli ; nous examinerons ce point dans le chapitre suivant.

Il est intéressant de faire quelques remarques sur le passage de l'ancienne équation (10) aux équations (18). Ce passage est un peu analogue au passage de l'équation des ondes lumineuses aux équations de Maxwell. En effet l'équation de propagation des ondes lumineuses (dans le vide) est :

$$\Delta u - \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = 0. \tag{23}$$

C'est une équation du second ordre où u désigne une grandeur caractérisant la perturbation lumineuse. Dans la théorie électromagnétique, u peut être l'une quelconque des six composantes des deux champs, champ électrique et champ magnétique, de l'onde lumineuse. On a donc six équations du type (23) pour les six grandeurs h_x, h_y, h_z, H_x, H_y et H_z.

Le passage de ces équations de propagation aux équations de Maxwell consiste précisément à remplacer les équations du second ordre par des équations du premier ordre simultanées liant les six grandeurs h_x ... H_z de telle façon que les six équations du second ordre du type (23) en soient la conséquence. C'est bien aussi la marche suivie par Dirac pour passer des équations du second ordre (15) aux équations du premier ordre (18).

Si l'onde de la Mécanique ondulatoire était une réalité physique au sens classique, on devrait s'attendre à ce que les quatre Ψ_i fussent les quatre composantes d'un vecteur d'espace-temps. Mais nous savons aujourd'hui que l'onde de la Mécanique ondulatoire n'est pas une réalité physique au sens classique : c'est une expression complexe servant seulement d'intermédiaire de calcul et permettant de former certaines expressions réelles ayant, elles, un sens physique telles que la densité de probabilité $\Psi\Psi^*$. Dans la théorie de Dirac, les 4 fonctions Ψ_i n'ont donc nullement besoin d'avoir le caractère de composantes de vecteur et nous verrons qu'effectivement elles ne l'ont pas. Mais il existe certaines com-

binaisons réelles de ces quantités qui ont un sens physique (de probabilité) et possèdent le caractère vectoriel. Nous apprendrons bientôt à les connaître.

4. ***Les équations de Dirac dans un champ électromagnétique.*** — Les équations (18) ne conviennent que pour le cas d'absence de champ. Que faut-il prendre comme équations de propagation si l'électron se déplace dans un champ défini par un potentiel scalaire V et un potentiel vecteur $\vec{A}$? Dirac résout la question en disant : il suffit de remplacer dans l'équation symbolique (12) les opérateurs p_i des formules (9) par les opérateurs P_i des formules (6).

Si l'on admet ce postulat de Dirac, les équations de propagation dans un champ électromagnétique s'écrivent symboliquement :

$$(P_4 + \alpha_1 P_1 + \alpha_2 P_2 + \alpha_3 P_3 + \alpha_4 m_0 c)\Psi = 0 \qquad (24)$$

ou explicitement :

$$\begin{aligned}
(P_4 + m_0 c)\Psi_1 + (P_1 + i P_2)\Psi_4 + P_3 \Psi_3 &= 0\\
(P_4 + m_0 c)\Psi_2 + (P_1 - i P_2)\Psi_3 - P_3 \Psi_4 &= 0\\
(P_4 - m_0 c)\Psi_3 + (P_1 + i P_2)\Psi_2 + P_3 \Psi_1 &= 0\\
(P_4 - m_0 c)\Psi_4 + (P_1 - i P_2)\Psi_1 - P_3 \Psi_2 &= 0
\end{aligned} \qquad (25)$$

Ce qu'il y a de vraiment remarquable, c'est que ces équations obtenues par des considérations générales tout à fait indépendantes des difficultés signalées dans la première partie du présent ouvrage, contiennent les propriétés de l'électron magnétique et tournant ! Nous allons commencer à nous en apercevoir en cherchant à former à partir des équations (25) les équations du second ordre qui généralisent les équations (15) dans le cas de la présence d'un champ électromagnétique.

Pour cela appliquons à l'équation symbolique (24) l'opérateur $P_4 - \left(\sum_1^3{}_i \alpha_i P_i + \alpha_4 m_0 c\right)$. Nous obtenons :

$$\left(P_4 - \sum_1^3{}_i \alpha_i P_i - \alpha_4 m_0 c\right)\left(P_4 + \sum_1^3{}_i \alpha_i P_i + \alpha_4 m_0 c\right)\Psi = 0 \qquad (26)$$

Si nous développons les opérateurs indiqués en tenant compte des relations (16), nous trouvons :

$$\Big[P_4^2 + \sum_1^3{}_i \alpha_i (P_4 P_i - P_i P_4) - \sum_1^3{}_i P_i^2$$
$$- \sum_{i \neq j} (\alpha_i \alpha_j P_i P_j + \alpha_j \alpha_i P_j P_i) - m_0^2 c^2\Big] \Psi = 0. \quad (27)$$

Rappelons-nous maintenant que les champs électromagnétiques sont reliés aux potentiels par les relations :

$$\vec{h} = -\operatorname{grad} V - \frac{1}{c}\frac{\partial \vec{A}}{\partial t}; \qquad \vec{H} = \operatorname{rot} \vec{A} \quad (28)$$

En tenant compte des définitions (6) et des conditions (16), nous obtenons après un calcul facile :

$$\sum_1^3{}_i \alpha_i (P_4 P_i - P_i P_4) = -\frac{h}{2\pi i}\cdot\frac{e}{c}(\alpha_1 h_x + \alpha_2 h_y + \alpha_3 h_z)$$
$$(29)$$
$$\sum_{i \neq j} (\alpha_i \alpha_j P_i P_j + \alpha_j \alpha_i P_j P_i) = -\frac{h}{2\pi i}\cdot\frac{e}{c}(\alpha_2 \alpha_3 H_x + \alpha_3 \alpha_1 H_y + \alpha_1 \alpha_2 H_z).$$

L'équation (27) devient ainsi :

$$\Big[P_4^2 - \sum_1^3{}_i P_i^2 - m_0^2 c_2 - \frac{e}{c}\cdot\frac{h}{2\pi i}(\alpha_1 h_x + \alpha_2 h_y + \alpha_3 h_z)$$
$$+ \frac{e}{c}\frac{h}{2\pi i}(\alpha_2 \alpha_3 H_x + \alpha_3 \alpha_1 H_y + \alpha_1 \alpha_2 H_z)\Big]\Psi = 0 \quad (30)$$

Si les trois premiers termes du crochet existaient seuls, on retomberait sur l'équation (5) qui serait ainsi valable pour chaque Ψ_i. L'élément nouveau introduit par la théorie de Dirac, c'est l'intervention des termes supplémentaires dans (30). Pour découvrir le sens de ces termes supplémentaires, regardons l'équation non relativiste (2) que nous pouvons écrire sous la forme :

$$\Big(2 m_0 c \cdot p_4 - \sum_1^3{}_i p_i^2 - 2 m_0 U\Big)\Psi = 0 \quad (31)$$

En comparant (31) avec (30), nous apercevons que nous pouvons considérer les termes supplémentaires en question comme des termes d'énergie potentielle mais à condition de les diviser par $-2m_0$. Seulement, en Mécanique de Dirac, le rôle joué par m_0 dans les théories antérieures est joué par $-\alpha_4 m_0$ comme nous le verrons plus tard [1]. Nous conviendrons donc de définir deux termes d'énergie potentielle en multipliant *en avant* les deux termes supplémentaires de (30) par $\frac{1}{2m_0\alpha_4} = \frac{\alpha_4}{2m_0}$. Compte tenu des relations de permutation entre les α_i, ceci nous conduit à poser :

$$\begin{aligned} U_e &= -\frac{eh}{4\pi m_0 c} i(\alpha_1\alpha_4 h_x + \alpha_2\alpha_4 h_y + \alpha_3\alpha_4 h_z) \\ U_m &= -\frac{eh}{4\pi m_0 c} i(\alpha_2\alpha_3\alpha_4 H_x + \alpha_3\alpha_1\alpha_4 H_y + \alpha_1\alpha_2\alpha_4 H_z) \end{aligned} \qquad (32)$$

Or il faut nous souvenir qu'un corps de moment électrique $\vec{\mathfrak{P}}$ placé dans un champ électrique $\vec{h}$ possède une énergie potentielle $-(\vec{\mathfrak{P}}\,.\,\vec{h})$ tandis qu'un corps de moment magnétique $\vec{\mathfrak{M}}$ placé dans un champ magnétique $\vec{H}$ possède une énergie potentielle $-(\vec{\mathfrak{M}}\,.\,\vec{H})$. Nous sommes donc amenés à attribuer à l'électron un moment magnétique de composantes :

$$\mathfrak{M}_x = \frac{eh}{4\pi m_0 c} i\alpha_2\alpha_3\alpha_4\,; \quad \mathfrak{M}_y = \frac{eh}{4\pi m_0 c} i\alpha_3\alpha_1\alpha_4\,;$$
$$\mathfrak{M}_z = \frac{eh}{4\pi m_0 c} i\alpha_1\alpha_2\alpha_4 \qquad (33)$$

et un moment électrique de composantes :

$$\mathfrak{P}_x = \frac{eh}{4\pi m_0 c} i\alpha_1\alpha_4\,; \quad \mathfrak{P}_y = \frac{eh}{4\pi m_0 c} i\alpha_2\alpha_4\,; \quad \mathfrak{P}_z = \frac{eh}{4\pi m_0 c} i\alpha_3\alpha_4 \qquad (34$$

(1) Voir chapitre XV, paragr. 4.

Pour parler avec plus de précision, les quantités (33) et (34) sont les opérateurs correspondant aux composantes des deux moments. Comme $\frac{eh}{4\pi m_0 c}$ est le magnéton de Bohr, on voit que les équations de Dirac attribuent automatiquement à l'électron un magnétisme propre relié au magnéton de Bohr. Nous voyons donc apparaître le magnétisme de l'électron et nous aurons bientôt l'occasion de préciser cette première indication. Les formules (34) nous montrent d'ailleurs que l'électron de Dirac possède aussi un moment électrique dont nous verrons plus tard la signification.

Les composantes des deux moments définies par (33) et (34) sont des opérateurs puisqu'elles s'expriment à l'aide des α_i. Ceci ne doit pas nous étonner puisque nous sommes habitués à voir dans la nouvelle Mécanique les grandeurs physiques céder la place à des opérateurs. Il est de plus, facile de vérifier que les opérateurs (33) et (34) sont hermitiques ; en effet les α_i étant hermitiques et anticommutant entre eux, les produits tels que $\alpha_1 \alpha_2$ ou $\alpha_1 \alpha_2 \alpha_4$ par exemple sont anti-hermitiques et, par suite, leur produit par i est hermitique. Les opérateurs (33) et (34) ont donc bien le caractère voulu pour représenter des grandeurs physiques.

Ainsi se trouve écartée une difficulté qui existait dans la présentation primitive de Dirac qui écrivait les formules (33) et (34) sans le facteur α_4 et obtenait ainsi des composantes de moment électrique non hermitiques.

CHAPITRE XI

Invariance relativiste des équations de Dirac

1. *Invariance de forme des équations de Dirac pour une transformation de Lorentz.*

LES équations de Dirac font jouer à l'axe des z un rôle particulier et, comme dans la théorie de Pauli, les fonctions d'onde serviront à répondre à des questions de probabilité où l'axe des z intervient. Si nous voulons poser les mêmes questions de probabilités pour un axe oz' distinct de oz, nous devrons écrire les équations de Dirac dans un système où oz' sera le troisième axe. Nous devons pouvoir écrire les équations de Dirac *sous la même forme* pour tous les systèmes d'axes, les quatre fonctions d'onde se transformant d'une certaine façon quand nous passons d'un système d'axes à l'autre.

Non seulement les équations de Dirac possèdent bien, comme les équations de Pauli, cette invariance de forme par rapport aux changements de coordonnées d'espace, mais elles sont aussi invariantes de forme pour toutes les transformations de Lorentz et satisfont ainsi au principe de Relativité.

On sait qu'une transformation de Lorentz du type le plus général peut toujours se décomposer en trois transforma-

tions successives. Partant du système primitif $xyzt$ de coordonnées d'espace temps, on opère :

1° Une rotation autour de oz définie par des formules du type :

$$x = x' \cos\alpha - y' \sin\alpha\,;\ y = y' \cos\alpha + x' \sin\alpha\,;\ z = z'\,;\ t = t' \tag{1}$$

2° Une rotation autour de l'axe oy perpendiculaire à oz définie par :

$$z = z' \cos\theta - x' \sin\theta\,;\ x = x' \cos\theta + z' \sin\theta\,;\ y = y'\,;\ t = t' \tag{2}$$

3° Une transformation *simple* de Lorentz, c'est-à-dire le passage d'un système $oxyzt$ à un système $o'x'y'z't'$ en mouvement relatif uniforme par rapport au premier, les axes z et z' glissant l'un sur l'autre et les autres axes restant respectivement parallèles ; on a alors les formules de transformation bien connues :

$$x = x'\,;\ y = y'\,;\ z = \frac{z' + \beta ct'}{\sqrt{1-\beta^2}}\,;\quad t = \frac{t' + \frac{\beta}{c} x'}{\sqrt{1-\beta^2}} \tag{3}$$

où βc est la vitesse du second système par rapport au premier. Si nous posons :

$$\frac{1}{\sqrt{1-\beta^2}} = ch\,\gamma \qquad \frac{\beta}{\sqrt{1-\beta^2}} = \sqrt{ch^2\gamma - 1} = sh\,\gamma \tag{4}$$

les formules (3) s'écrivent :

$$x = x'\,;\ y = y'\,;\quad z = z'\,ch\,\gamma + ct' sh\,\gamma\,;$$
$$ct = ct'\,ch\,\gamma + z'\,sh\,\gamma\,. \tag{5}$$

En combinant trois formations des types 1°, 2° et 3°, on peut effectuer une transformation de Lorentz du type le plus général

L'invariance de forme des équations de Dirac pour une transformation générale de Lorentz sera prouvée si l'on parvient à résoudre le problème suivant : sachant que dans un

système d'axes galiléens $xyzt$, les équations de Dirac :

$$(P_4 + \alpha_1 P_1 + \alpha_2 P_2 + \alpha_3 P_3 + \alpha_4 m_0 c) \Psi = 0 \qquad (6)$$

admettent pour solutions les fonctions $\Psi_1(xyzt)$, $\Psi_2(xyzt)$, $\Psi_3(xyzt)$, $\Psi_4(xyzt)$, montrer que dans un autre système d'axes galiléens $x'\, y'\, z'\, t'$ les équations de Dirac :

$$(P_4' + \alpha_1 P_1' + \alpha_2 P_2' + \alpha_3 P_3' + \alpha_4 m_0 c) \Psi' = 0 \qquad (7)$$

où l'on a :

$$P_1' = -\frac{h}{2\pi i}\frac{\partial}{\partial x'} + \frac{e}{c} A_x', \ \ldots\ldots \quad P_4' = \frac{h}{2\pi i}\frac{1}{c}\frac{\partial}{\partial t'} + \frac{e}{c} V' \qquad (8)$$

admettent pour solutions des fonctions $\Psi_1'(x'\,y'\,z'\,t')$, $\Psi_2'(x'\,y'\,z'\,t')$, $\Psi_3'(x'\,y'\,z'\,t')$ et $\Psi_4'(x'\,y'\,z'\,t')$, s'exprimant linéairement en fonction de $\Psi_1 \ldots \Psi_4$ par des formules où figurent les paramètres qui définissent le passage du système $oxyzt$ au système $o'\,x'\,y'\,z'\,t'$.

Il suffit de résoudre le problème ainsi posé par les trois transformations 1°, 2° et 3° indiquées plus haut puisque toute transformation de Lorentz peut se décomposer en transformations de ces trois types.

1° Rotation autour de oz :

Partons des équations :

$$\begin{aligned}
&(P_4 + m_0 c)\Psi_1 + (P_1 + i P_2)\Psi_4 + P_3 \Psi_3 = 0\\
&(P_4 + m_0 c)\Psi_2 + (P_1 - i P_2)\Psi_3 - P_3 \Psi_4 = 0\\
&(P_4 - m_0 c)\Psi_3 + (P_1 + i P_2)\Psi_2 + P_3 \Psi_1 = 0\\
&(P_4 - m_0 c)\Psi_4 + (P_1 - i P_2)\Psi_1 - P_3 \Psi_2 = 0 .
\end{aligned} \qquad (9)$$

Comme le changement de variables s'exprime par les formules (1), on a :

$$P_1 = P_1' \cos\alpha - P_2' \sin\alpha \; ; \; P_2 = P_2' \cos\alpha + P_1' \sin\alpha$$
$$P_3 = P_3' \; ; \; P_4 = P_4' \qquad (1')$$

et en substituant dans (9) on trouve aisément :

$$\begin{aligned}
&(P_4' + m_0 c)\Psi_1 + (P_1' e^{i\alpha} + i P_2' e^{i\alpha})\Psi_4 + P_3' \Psi_3 = 0\\
&(P_4' + m_0 c)\Psi_2 + (P_1' e^{-i\alpha} - i P_2' e^{-i\alpha})\Psi_3 - P_3' \Psi_4 = 0\\
&(P_4' - m_0 c)\Psi_3 + (P_1' e^{i\alpha} + i P_2' e^{i\alpha})\Psi_2 + P_3' \Psi_1 = 0\\
&(P_4' - m_0 c)\Psi_4 + (P_1' e^{-i\alpha} - i P_2' e^{-i\alpha})\Psi_1 - P_3' \Psi_2 = 0
\end{aligned} \qquad (10)$$

Dans ces équations les Ψ_i sont supposées exprimées en fonction des variables primées à l'aide des relations de transformation (1). Multiplions la première et la troisième équations (10) par $e^{-i\frac{\alpha}{2}}$, la deuxième et la quatrième par $e^{i\frac{\alpha}{2}}$ nous obtenons :

$$
\begin{aligned}
&(P_4' + m_0c)\,\Psi_1\,e^{-i\frac{\alpha}{2}} + (P_1' + i\,P_2')\,\Psi_4\,e^{i\frac{\alpha}{2}} + P_3'\,\Psi_3\,e^{-i\frac{\alpha}{2}} = 0 \\
&(P_4' + m_0c)\,\Psi_2\,e^{i\frac{\alpha}{2}} + (P_1' - i\,P_2')\,\Psi_3\,e^{-i\frac{\alpha}{2}} - P_3'\,\Psi_4\,e^{i\frac{\alpha}{2}} = 0 \\
&(P_4' - m_0c)\,\Psi_3\,e^{-i\frac{\alpha}{2}} + (P_1' + i\,P_2')\,\Psi_1 e^{-i\frac{\alpha}{2}} + P_3'\,\Psi_1\,e^{-i\frac{\alpha}{2}} = 0 \\
&(P_4' - m_0c)\,\Psi_4\,e^{i\frac{\alpha}{2}} + (P_1' - i\,P_2')\,\Psi_2 e^{i\frac{\alpha}{2}} - P_3'\,\Psi_2\,e^{i\frac{\alpha}{2}} = 0.
\end{aligned}
\tag{11}
$$

Le système (11) nous montre que les fonctions Ψ_i' sont reliées aux fonction Ψ_i par les formules simples :

$$
\begin{aligned}
&\Psi_1'(x'\,y'\,z'\,t') = \Psi_1(x'\,y'\,z'\,t')\,e^{-i\frac{\alpha}{2}}\,; \\
&\Psi_2'(x'\,y'\,z'\,t') = \Psi_2(x'\,y'\,z'\,t')\,e^{i\frac{\alpha}{2}} \\
&\Psi_3'(x'\,y'\,z'\,t') = \Psi_3(x'\,y'\,z'\,t')\,e^{-i\frac{\alpha}{2}}\,; \\
&\Psi_4'(x'\,y'\,z'\,t') = \Psi_4(x'\,y'\,z'\,t')\,e^{i\frac{\alpha}{2}}
\end{aligned}
\tag{12}
$$

et la démonstration demandée se trouve faite en ce cas.

2° Rotation autour de oy :

La transformation des coordonnées étant donnée par (2), la transformation des fonctions d'onde de Dirac est donnée par les formules :

$$
\begin{aligned}
&\Psi_1'(x'\,y'\,z'\,t') = \Psi_1(x'\,y'\,z'\,t')\cos\frac{\theta}{2} + \Psi_2(x'\,y'\,z'\,t')\sin\frac{\theta}{2} \\
&\Psi_2'(x'\,y'\,z'\,t') = \Psi_2(x'\,y'\,z'\,t')\cos\frac{\theta}{2} - \Psi_1(x'\,y'\,z'\,t')\sin\frac{\theta}{2} \\
&\Psi_3'(x'\,y'\,z'\,t') = \Psi_3(x'\,y'\,z'\,t')\cos\frac{\theta}{2} + \Psi_4(x'\,y'\,z'\,t')\sin\frac{\theta}{2} \\
&\Psi_4'(x'\,y'\,z'\,t') = \Psi_4(x'\,y'\,z'\,t')\cos\frac{\theta}{2} - \Psi_3(x'\,y'\,z'\,t')\sin\frac{\theta}{2}
\end{aligned}
\tag{13}
$$

On le démontre par un raisonnement analogue à celui qui a été fait pour le cas 1°.

3° Transformation simple de Lorentz :

La transformation des coordonnées étant exprimée par les formules (5), celle des fonctions de Dirac est donnée par :

$$\begin{aligned}
\Psi_1'(x'\,y'\,z'\,t') &= \Psi_1(x'\,y'\,z'\,t')\,ch\frac{\gamma}{2} + \Psi_3(x'\,y'\,z'\,t')\,sh\frac{\gamma}{2}\\
\Psi_2'(x'\,y'\,z'\,t') &= \Psi_2(x'\,y'\,z'\,t')\,ch\frac{\gamma}{2} - \Psi_4(x'\,y'\,z'\,t')\,sh\frac{\gamma}{2}\\
\Psi_3'(x'\,y'\,z'\,t') &= \Psi_3(x'\,y'\,z'\,t')\,ch\frac{\gamma}{2} + \Psi_1(x'\,y'\,z'\,t')\,sh\frac{\gamma}{2}\\
\Psi_4'(x'\,y'\,z'\,t') &= \Psi_4(x'\,y'\,z'\,t')\,ch\frac{\gamma}{2} - \Psi_2(x'\,y'\,z'\,t')\,sh\frac{\gamma}{2}
\end{aligned} \tag{14}$$

Ceci se démontre encore de la même façon et l'invariance de forme des équations de Dirac pour la transformation de Lorentz la plus générale est ainsi prouvée.

On voit sur les formules (12), (13) et (14) que les Ψ_i ne se transforment pas comme les coordonnées ; elles n'ont donc pas, comme nous l'avions déjà dit, le caractère de composantes d'un vecteur d'espace-temps. Mais nous apprendrons bientôt à former avec les Ψ_i certaines expressions qui auront, elles, le caractère vectoriel ou tensoriel et auront une signification physique.

2. ***Démonstration plus synthétique de l'invariance relativiste.*** — Nous allons indiquer une démonstration plus synthétique de l'invariance des équations de Dirac ; cette démonstration développée par M. von Neumann est une généralisation de celle qu'avait employée M. Pauli pour démontrer l'invariance de ses équations par rapport au changement de coordonnées rectangulaires dans l'espace.

Prenons l'équation de Dirac sous sa forme symbolique :

$$\left(P_4 + \sum_1^3{}_i \alpha_i\,P_i + \alpha_4\,m_0 c\right)\Psi = 0 \tag{15}$$

et appliquons-lui d'abord l'opérateur α_4. Nous obtenons ($\alpha_4^2 = 1$) :

$$\left(\alpha_4 P_4 + \sum_1^3{}_i \alpha_4 \alpha_i P_i + m_0 c\right) \Psi = 0 \tag{16}$$

Nous allons prendre maintenant, à la place des variables x, y, z, t, les variables d'univers de Minkowski :

$$x_1 = x \quad x_2 = y \quad x_3 = z \quad x_4 = ict \tag{17}$$

et nous définirons les opérateurs correspondant :

$$\pi_1 = -\frac{h}{2\pi i}\frac{\partial}{\partial x_1} + \frac{e A_1}{c} = P_1 \,; \quad \pi_2 = P_2 \,; \quad \pi_3 = P_3 \,;$$
$$\pi_4 = -\frac{h}{2\pi i}\frac{\partial}{\partial x_4} + \frac{e}{c} i V = -\frac{P_4}{i} \tag{18}$$

L'équation (16) s'écrit alors

$$\left(\alpha_4 \pi_4 + \sum_1^3{}_i i\, \alpha_4 \alpha_i \pi_i + i\, m_0 c\right) \Psi = 0 \tag{19}$$

Posons alors (von Neumann) :

$$\gamma_1 = i\, \alpha_4 \alpha_1 \quad \gamma_2 = i\, \alpha_4 \alpha_2 \quad \gamma_3 = i\, \alpha_4 \alpha_3 \quad \gamma_4 = \alpha_4 \tag{20}$$

Il est facile de vérifier que l'on a :

$$\gamma_i^2 = 1 \quad \gamma_i \gamma_j + \gamma_j \gamma_i = 0 \quad (i \neq j) \tag{21}$$

ce que l'on peut résumer en écrivant :

$$\gamma_i \gamma_j + \gamma_j \gamma_i = 2\, \delta_{ij} \,.\, 1 \tag{22}$$

1 étant la matrice unité. On vérifie aussi que les γ_i sont hermitiques.

Avec ces notations, l'équation (19) prend la forme condensée souvent utile :

$$\left(\sum_1^4{}_i \gamma_i \pi_i + i\, m_0 c\right) \Psi = 0 \tag{23}$$

Supposons maintenant que nous fassions un changement

d'axes galiléens, une transformation générale de Lorentz. Il est bien connu en théorie de Relativité qu'une transformation générale de Lorentz équivaut à une rotation des axes dans l'univers de Minkowski. Les nouvelles variables x_i' après la transformation seront donc reliées aux anciennes variables x_i par les formules :

$$x_i = \sum_j o_{ij} x'_j \tag{24}$$

o étant une matrice à 4 rangs et 4 colonnes [1] satisfaisant à la relation d'orthogonalité :

$$\sum_i o_{ik} o_{ij} = \delta_{kj} \tag{25}$$

Or il est visible que les π_i se transforment comme les x_i c'est-à-dire que :

$$\pi_i = \sum_j o_{ij} \pi'_j \tag{26}$$

Après le changement de coordonnées, l'équation (23) s'écrira donc :

$$\left(\sum_1^4{}_i \gamma_i \sum_1^4{}_j o_{ij} \pi'_j + i\, m_0\, c\right) \Psi = 0 \tag{27}$$

où les fonctions Ψ devront être exprimées à l'aide des nouvelles variables x_i'. Si nous posons :

$$\gamma'_j = \sum_1^4{}_i o_{ij} \gamma_i \tag{28}$$

nous pourrons remplacer (27) par :

$$\left(\sum_1^4{}_j \gamma_j' \pi_j' + i\, m_0\, c\right) \Psi = 0 \tag{29}$$

L'équation (28) exprime les matrices γ_j' en fonction des matrices γ_i. Si l'on veut préciser son sens, on écrira :

$$\gamma'_{j,mn} = \sum_1^4{}_i o_{ij} \gamma_{i,mn} \tag{30}$$

[1] La matrice 0 n'est pas réelle ; ceux de ses éléments qui contiennent une fois l'indice 4 sont purement imaginaires

où $\gamma_{i,mn}$ par exemple désigne l'élément d'indices m, n de la matrice γ_i.

Comme l'on a :

$$\gamma'^*_{j,nm} = \sum_{1}^{4}{}_i\, o^*_{ij}\gamma^*_{i,nm} = \sum_{1}^{4}{}_i\, o^*_{ij}\gamma_{i,mn} \neq \gamma'_{j,mn} \qquad (31)$$

en raison de l'hermiticité des γ_i, on voit qu'en général les γ_i' ne sont pas hermitiques parce que tous les o_{ij} ne sont pas réels.

Il est facile de vérifier que les γ_i' satisfont aux conditions de la forme (22). On a en effet :

$$\begin{aligned} \gamma_i'\gamma_j' + \gamma_j'\gamma_i' &= \sum_k o_{ki}\gamma_k \cdot \sum_l o_{lj}\gamma_l + \sum_l o_{lj}\gamma_l \cdot \sum_k o_{ki}\gamma_k \\ &= \sum_k\sum_l o_{ki}\, o_{lj}(\gamma_k\gamma_l + \gamma_l\gamma_k) \qquad (32) \\ &= \sum_k\sum_l o_{ki}\, o_{lj} \,.\, 2\,\delta_{kl} = 2\sum_k o_{ki}\, o_{kj} = 2\,\delta_{ij} \end{aligned}$$

Le point essentiel du raisonnement est qu'il existe une matrice Λ à 4 lignes et à 4 colonnes telle que l'on ait :

$$\gamma_i' = \Lambda^{-1}\gamma_i\Lambda \qquad (i = 1, 2, 3, 4,) \qquad (33)$$

Si la relation (33) est vraie, le fait que les γ_i satisfont aux relations (22) entraîne que les γ_i' satisfont aux relations analogues :

$$\gamma_i'\,\gamma_j' + \gamma_j'\,\gamma_i' = 2\,\delta_{ij} \qquad (34)$$

L'équation (33) pouvant s'écrire :

$$\Lambda\gamma_i' = \gamma_i\Lambda \qquad (i = 1, 2, 3, 4) \qquad (35)$$

pour démontrer *l'existence* de la matrice Λ, il faudrait montrer qu'il existe 16 grandeurs Λ_{kl} satisfaisant aux 64 équations :

$$\sum_{1}^{4}{}_l\, \Lambda_{kl}\gamma'_{i,lm} = \sum_{1}^{4}{}_l\, \gamma_{i,kl}\Lambda_{lm} \qquad (36)$$

où les indices k, m, i peuvent prendre les valeurs 1, 2, 3, 4. L'existence de la matrice Λ n'est donc nullement évidente

a priori. Nous allons cependant admettre provisoirement l'existence de cette matrice, quitte à justifier ensuite cette hypothèse *a posteriori*.

Avant d'aller plus loin, remarquons que la matrice Λ ne peut être unitaire en général. En effet, si elle était unitaire. on aurait :

$$\Lambda^{+} = \Lambda^{-1} ; \qquad \Lambda = (\Lambda^{-1})^{+} \tag{37}$$

et comme on tire de (33) en prenant l'équation adjointe :

$$\gamma_i'^{+} = (\Lambda^{-1} \gamma_i \Lambda)^{+} = \Lambda^{+} \gamma_i^{+} (\Lambda^{-1})^{+} \tag{38}$$

on aurait en vertu de (37) et de l'hermiticité de γ_i :

$$\gamma_i'^{+} = \Lambda^{-1} \gamma_i \Lambda = \gamma_i' \tag{39}$$

Il en résulterait que γ_i' serait hermitique, ce qui, nous l'avons vu, n'a pas lieu en général.

Reprenons maintenant l'équation (29) en y introduisant la relation (33). Il vient :

$$\left(\sum_{1}^{4}{}_{j} \Lambda^{-1} \gamma_j \Lambda \pi_j' + i m_0 c\right) \Psi' = 0 \tag{40}$$

Multiplions en avant par Λ et remarquons que la matrice Λ correspondant à une opération effectuée sur l'indice de Dirac, permute avec π_j'. Nous obtenons :

$$\left(\sum_{1}^{4}{}_{j} \gamma_j \pi_j' + i m_0 c\right) \Lambda \Psi' = 0 \tag{41}$$

La formule (41) exprime le théorème suivant :

Théorème : Quand on fait une transformation de Lorentz, on peut conserver aux équations de Dirac la même forme, mais les nouvelles fonctions d'onde Ψ_i', sont reliées aux anciennes, Ψ_i, par une transformation linéaire :

$$\Psi_i'(x', y', z', t') = \Lambda \Psi_i(x', y', z', t') = \sum_{1}^{4}{}_{k} \Lambda_{ik} \Psi_k(x', y', z', t') \tag{42}$$

C'est précisément là le résultat déjà établi au paragraphe 1.

Nous pouvons maintenant justifier l'hypothèse de l'existence de la matrice Λ. Nous avons en effet appris au paragraphe 1 à calculer la transformation linéaire que subit chaque fonction Ψ_i pour chacune des trois sortes de changement de coordonnées en lesquelles on peut décomposer toute transformation générale de Lorentz. Nous savons donc, du moins en principe, comment calculer la transformation linéaire des Ψ_i qui correspond à une transformation de Lorentz quelconque, c'est-à-dire déterminer les éléments de la matrice Λ telle que :

$$\Psi_i'(x'y'z't') = \Lambda \Psi_i(x', y', z', t) \qquad (43)$$

Or cette matrice Λ doit d'après (42) être la même que celle qui figure dans (33) et puisque nous savons la calculer, nous sommes sûrs qu'elle existe.

En se reportant aux formules (12), (13) et (14), on peut écrire immédiatement l'expression des matrices Λ qui correspondent aux cas simples 1°, 2° et 3° du paragraphe précédent. Considérons par exemple le cas 3° : transformation simple de Lorentz. Les formules (14) nous montrent que la matrice Λ a alors la forme suivante :

$$\Lambda = \begin{vmatrix} ch\frac{\gamma}{2} & o & sh\frac{\gamma}{2} & o \\ o & ch\frac{\gamma}{2} & o & -sh\frac{\gamma}{2} \\ sh\frac{\gamma}{2} & o & ch\frac{\gamma}{2} & o \\ o & -sh\frac{\gamma}{2} & o & ch\frac{\gamma}{2} \end{vmatrix} \qquad (44)$$

Il est facile de vérifier que cette matrice n'est pas unitaire ($\Lambda^+ \neq \Lambda^{-1}$). Il en résulte que $\sum_{1}^{4}{}_i \Psi_i^* \Psi_i$ n'est pas égal à $\sum_{1}^{4}{}_i \Psi_i'^* \Psi_i'$. La densité de probabilité change donc de valeur pour une transformation simple de Lorentz. Nous verrons en effet que cette densité n'est pas un invariant, mais la composante de temps d'un vecteur d'espace-temps.

3. ***Invariance électromagnétique des équations de Dirac.*** — A côté de l'invariance relativiste, les équations de Dirac présentent une autre espèce d'invariance que je nommerai l'invariance électromagnétique (c'est l'invariance de jauge, *Eichinvarianz*, de auteurs allemands). Expliquons en quoi elle consiste.

Le champ électrique $\vec{h}$ et le champ magnétique $\vec{H}$ étant définis en fonction des potentiels par les formules :

$$\vec{h} = -\operatorname{grad} V - \frac{1}{c}\frac{\partial \vec{A}}{\partial t} \qquad \vec{H} = \operatorname{rot} \vec{A} \tag{45}$$

il est visible que si l'on remplace V et $\vec{A}$ par :

$$V' = V - \frac{1}{c}\frac{\partial \Phi}{\partial t} \qquad \vec{A'} = \vec{A} + \operatorname{grad} \Phi \tag{46}$$

Φ étant une fonction quelconque de $xyzt$, on ne modifiera aucunement les champs car on a :

$$\operatorname{rot} A' = \operatorname{rot} A \qquad -\operatorname{grad} V' - \frac{1}{c}\frac{\partial \vec{A'}}{\partial t} = -\operatorname{grad} V - \frac{1}{c}\frac{\partial \vec{A}}{\partial t} \tag{47}$$

Puisque ce sont les champs qui expriment les actions dynamiques et puisqu'ils sont insensibles à une transformation des potentiels de la forme (46), nous devons nous attendre à ce que les équations de Dirac soient invariantes par rapport à ces transformations (46). C'est là l'invariance électromagnétique en question.

Ecrivons l'équation symbolique de Dirac :

$$\left[\left(\frac{h}{2\pi i}\frac{1}{c}\frac{\partial}{\partial t} + \frac{eV}{c}\right) + \sum_1^3{}_j \alpha_j\left(-\frac{h}{2\pi i}\frac{\partial}{\partial x_j} + \frac{e}{c}A_j\right) + \alpha_4 m_0 c\right]\Psi = 0 \tag{48}$$

puis supposons que nous fassions subir aux potentiels la transformation (46). En remplaçant V et A_j en fonction de

V' et A_j', on obtient à la place de (48) :

$$\left[\left(\frac{h}{2\pi i}\frac{1}{c}\frac{\partial}{\partial t}+\frac{eV'}{c}+\frac{e}{c^2}\frac{\partial\Phi}{\partial t}\right)+\sum_{j}^{3}{}_{1}\alpha_j\left(-\frac{h}{2\pi i}\frac{\partial}{\partial x_j}+\frac{e}{c}A_j'-\frac{e}{c}\frac{\partial\Phi}{\partial x_j}\right)+\alpha_4 m_0 c\right]\Psi=0 \qquad (49)$$

Il est aisé de vérifier que si l'on pose : $\Psi'=\Psi\, e^{-\frac{2\pi i}{h}\frac{e}{c}\Phi}$, on retrouve une équation ayant la même forme que (48), savoir :

$$\left[\left(\frac{h}{2\pi i}\frac{1}{c}\frac{\partial}{\partial t}+\frac{e}{c}V'\right)+\sum_{j}^{3}{}_{1}\alpha_j\left(-\frac{h}{2\pi i}\frac{\partial}{\partial x_j}+\frac{e}{c}A_j'\right)+\alpha_4 m_0 c\right]\Psi'=0 \qquad (50)$$

Le passage des potentiels V et $\vec{A}$ aux potentiels V' et $\vec{A'}$ ne modifie donc pas la forme des équations de Dirac mais chacun des Ψ_k est multiplié par $e^{-\frac{2\pi i}{h}\frac{e}{c}\Phi}$.

Mais nous savons que ce ne sont pas les Ψ_k mais certaines combinaisons des Ψ_k qui ont un sens physique ; or on pourra vérifier que toutes les combinaisons de ce genre que nous apprendrons bientôt à connaître ne sont pas modifiées quand on y remplace Ψ_k par $\Psi_k\, e^{-\frac{2\pi i}{h}\frac{e}{c}\Phi}$; ceci est immédiatement visible par exemple pour la densité de probabilité $\sum_{k=1}^{4}\Psi_k^*\Psi_k$. Cette insensibilité des quantités ayant un sens physique par rapport aux transformations (46) constitue l'invariance électromagnétique des équations de Dirac.

CHAPITRE XII

Densité et courant dans la théorie de Dirac. — Ondes planes

1. *Expressions de la densité et du courant de probabilité*

DANS la théorie de Dirac, nous devons chercher à transposer les idées de la Mécanique ondulatoire à une fonction Ψ. En particulier nous devons chercher à former une densité ρ de probabilité de présence et à définir un courant $\overrightarrow{\rho u}$ de cette probabilité. Les expressions $-e\rho$ et $-e\overrightarrow{\rho u}$ donneront ici encore la densité électrique moyenne et le courant électrique moyen à l'aide desquels se calculeront les rayonnements moyens émis par des ensembles d'électrons.

Pour trouver la forme de ρ et de $\overrightarrow{\rho u}$, nous devons toujours être guidés par l'idée que la probabilité totale de présence doit rester constante (égale à 1) et que par suite l'équation de continuité $\frac{\partial \rho}{\partial t} + \mathrm{div}(\overrightarrow{\rho u}) = 0$ doit être une conséquence des équations de propagation.

Ecrivons les 4 équations de Dirac et les 4 équations conjuguées :

$$
\begin{aligned}
&(P_4 + m_0 c)\,\Psi_1 + (P_1 + i\,P_2)\,\Psi_4 + P_3\,\Psi_3 = 0\\
&(P_4 + m_0 c)\,\Psi_2 + (P_1 - i\,P_2)\,\Psi_3 - P_3\,\Psi_4 = 0\\
&(P_4 - m_0 c)\,\Psi_3 + (P_1 + i\,P_2)\,\Psi_2 + P_3\,\Psi_1 = 0\\
&(P_4 - m_0 c)\,\Psi_4 + (P_1 - i\,P_2)\,\Psi_1 - P_3\,\Psi_2 = 0
\end{aligned}
\qquad (1)
$$

et :

$$
\begin{aligned}
&(P_4^* + m_0 c)\,\Psi_1^* + (P_1^* - i\,P_2^*)\,\Psi_4^* + P_3^*\,\Psi_3^* = 0\\
&(P_4^* + m_0 c)\,\Psi_2^* + (P_1^* + i\,P_2^*)\,\Psi_3^* - P_3^*\,\Psi_4^* = 0\\
&(P_4^* - m_0 c)\,\Psi_3^* + (P_1^* - i\,P_2^*)\,\Psi_2^* + P_3^*\,\Psi_1^* = 0\\
&(P_4^* - m_0 c)\,\Psi_4^* + (P_1^* + i\,P_2^*)\,\Psi_1^* - P_3^*\,\Psi_2^* = 0
\end{aligned}
\qquad (2)
$$

Mutiplions les équations (1) respectivement par Ψ_1^*, Ψ_2^*, Ψ_3^*, Ψ_4^* et les équations (2) respectivement par Ψ_1, Ψ_2, Ψ_3, Ψ_4. Faisons ensuite la somme des équations (1) et retranchons-en la somme des équations (2). Les termes en $m_0 c$ s'éliminent. On trouve des termes de la forme :

$$
\Psi_i^*\,P_4\,\Psi_i - \Psi_i\,P_4^*\,\Psi_i^* = \Psi_i^*\left(\frac{h}{2\pi i}\,\frac{1}{c}\,\frac{\partial}{\partial t} + \frac{e}{c}\,V\right)\Psi_i -
$$
$$
- \Psi_i^*\left(-\frac{h}{2\pi i}\,\frac{1}{c}\,\frac{\partial}{\partial t} + \frac{e}{c}\,V\right)\Psi_i^* = \frac{h}{2\pi i}\,\frac{1}{c}\,\frac{\partial}{\partial t}\,(\Psi_i^*\,\Psi_i)
$$

et aussi des termes de la forme :

$$
\Psi_i^*\,P_1\,\Psi_i - \Psi_i\,P_1^*\,\Psi_i^* = \Psi_i^*\left(-\frac{h}{2\pi i}\,\frac{\partial}{\partial x} + \frac{e}{c}\,A_x\right)\Psi_i -
$$
$$
-\Psi_i^*\left(\frac{h}{2\pi i}\,\frac{\partial}{\partial x} + \frac{e}{c}\,A_x\right)\Psi_i = -\frac{h}{2\pi i}\,\frac{\partial}{\partial x}\,(\Psi_i^*\,\Psi_i)
$$

Finalement, après multiplication par $\dfrac{2\pi i c}{h}$, on trouve :

$$
\frac{\partial}{\partial t}\left(\Psi_1^*\Psi_1 + \Psi_2^*\Psi_2 + \Psi_3^*\Psi_3 + \Psi_4^*\Psi_4\right) + \frac{\partial}{\partial x}\Big[-c\,(\Psi_1^*\Psi_4 +
$$
$$
\Psi_2^*\Psi_3 + \Psi_3^*\Psi_2 + \Psi_4^*\Psi_1)\Big] + \frac{\partial}{\partial y}\Big[-c\,(i\Psi_1^*\Psi_4 - i\Psi_2^*\Psi_3 +
$$
$$
i\,\Psi_3^*\Psi_2 - i\Psi_4^*\Psi_1)\Big] + \frac{\partial}{\partial z}\Big[-c\,(\Psi_1^*\Psi_3 - \Psi_2^*\Psi_4 + \Psi_3^*\Psi_1 -
$$
$$
\Psi_4^*\Psi_2)\Big] = 0 \qquad (3)
$$

L'équation (3) est équivalente à l'équation de continuité si l'on pose :

$$
\begin{aligned}
\rho &= \Psi_1^* \Psi_1 + \Psi_2^* \Psi_2 + \Psi_3^* \Psi_3 + \Psi_4^* \Psi_4) \\
\rho\, u_x &= -c\,(\Psi_1^* \Psi_4 + \Psi_2^* \Psi_3 + \Psi_3^* \Psi_2 + \Psi_4^* \Psi_1) \\
\rho\, u_y &= -c\,(i\,\Psi_1^* \Psi_4 - i\,\Psi_2^* \Psi_3 + i\,\Psi_3^* \Psi_2 - i\,\Psi_4^* \Psi_1) \\
\rho\, u_z &= -c\,(\Psi_1^* \Psi_3 - \Psi_2^* \Psi_4 + \Psi_3^* \Psi_1 - \Psi_4^* \Psi_2).
\end{aligned} \tag{4}
$$

On vérifie sans peine que les expressions (4) sont réelles car elles sont égales à leurs conjuguées. On peut donc les considérer comme définissant la densité de probabilité de présence et les composantes du courant correspondant.

L'expression de ρ a bien la forme postulée par Dirac. Elle montre d'ailleurs que l'on doit normaliser les Ψ, en écrivant :

$$
\iiint_{-\infty}^{+\infty} (\Psi_1^* \Psi_1 + \Psi_2^* \Psi_2 + \Psi_3^* \Psi_3 + \Psi_4^* \Psi_4)\, d\tau = 1 \tag{5}
$$

et que si cette normalisation est réalisée à une époque quelconque, elle restera ensuite toujours réalisée.

On peut donner une forme condensée et élégante aux formules (4) en se servant des trois matrices $\alpha_1 \alpha_2 \alpha_3$ et de la matrice 1 (à 4 rangs et 4 colonnes). Ecrivons ces matrices :

$$
\alpha_1 = \begin{vmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{vmatrix}; \quad
\alpha_2 = \begin{vmatrix} 0 & 0 & 0 & +i \\ 0 & 0 & -i & 0 \\ 0 & +i & 0 & 0 \\ -i & 0 & 0 & 0 \end{vmatrix}; \quad
\alpha_3 = \begin{vmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{vmatrix}
$$

$$
1 = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \tag{6}
$$

et comparons-les aux formules (4). Nous voyons aussitôt que celles-ci peuvent s'écrire :

$$
\begin{aligned}
\rho &= \sum_{1}^{4}{}_i\, \Psi_i^* \,.\, 1\, \Psi_i\,; \qquad \rho\, u_x = -c \sum_{1}^{4}{}_i\, \Psi_i^* \,.\, \alpha_1 \Psi_i\,; \\
\rho\, u_y &= -c \sum_{1}^{4}{}_i\, \Psi_i^* \,.\, \alpha_2 \Psi_i\,; \qquad \rho\, u_z = -c \sum_{1}^{4}{}_i\, \Psi_i^* \,.\, \alpha_3 \Psi_i.
\end{aligned} \tag{7}
$$

Naturellement la densité électrique moyenne δ et les com-

posantes du courant moyen j_x j_y et j_z s'obtiennent en multipliant ρ, ρu_x, etc. par $-e$. On a :

$$\delta = -e\rho\,;\; j_x = -e\rho u_x\,;\; j_y = -e\rho v_y\,;\; j_z = -e\rho v_z \qquad (8)$$

2. ***Caractère vectoriel de la densité et du courant.*** — Les quantités (4), qui ont un sens physique, présentent les caractères de composantes d'un vecteur d'espace-temps, ρ étant la composante de temps. Pour le démontrer, il suffit par exemple de vérifier que pour chacune des trois transformations envisagées au premier paragraphe du chapitre précédent, les quantités (4) se transforment comme les coordonnées, d'où il résulte qu'il en est de même pour les transformations de Lorentz les plus générales.

Indiquons la marche de la vérification pour une transformation du type 1° : rotation autour de oz. On a alors entre les anciennes et nouvelles variables les relations :

$$x = x'\cos\alpha - y'\sin\alpha \quad y = y'\cos\alpha + x'\sin\alpha \quad z = z' \quad t = t' \qquad (9)$$

d'où aussi :

$$x' = x\cos\alpha + y\sin\alpha \quad y' = y\cos\alpha - x\sin\alpha \quad z' = z \quad t' = t \quad (10)$$

Nous avons vu que les fonctions Ψ_i se transforment alors de la manière suivante :

$$\Psi_1' = \Psi_1\, e^{-i\frac{\alpha}{2}}\,;\; \Psi_2' = \Psi_2\, e^{+i\frac{\alpha}{2}}\,;$$
$$\Psi_3' = \Psi_3\, e^{-i\frac{\alpha}{2}}\,;\; \Psi_4' = \Psi_4\, e^{+i\frac{\alpha}{2}}\,. \quad (11)$$

Il est alors évident que l'on a :

$$\rho' = \sum_1^4{}_i \Psi_i'^* \Psi_i' = \sum_1^4{}_i \Psi_i^* \Psi_i = \rho. \qquad (12)$$

La densité ρ reste invariante dans la transformation comme la variable t. On voit tout aussi facilement que l'on a : $\rho' u_z' = \rho u_z$; la composante z du courant reste invariante comme la variable z. Pour la composante x, on a :

$$\begin{aligned}\rho' u_x' &= -c\,(\Psi_1'^* \Psi_4' + \Psi_2'^* \Psi_3' + \Psi_3'^* \Psi_2' + \Psi_4'^* \Psi_1') \\ &= -c\,(\Psi_1^* \Psi_4 e^{i\alpha} + \Psi_2^* \Psi_3 e^{-i\alpha} \\ &\qquad + \Psi_3^* \Psi_2 e^{i\alpha} + \Psi_4^* \Psi_1 e^{-i\alpha}) \\ &= -c\,(\Psi_1^* \Psi_4 + \Psi_2^* \Psi_3 + \Psi_3^* \Psi_2 + \Psi_4^* \Psi_1) \cos\alpha \\ &\quad - c\,(i\,\Psi_1^* \Psi_4 - i\,\Psi_2^* \Psi_3 + i\,\Psi_3^* \Psi_2 - i\,\Psi_4^* \Psi_1) \sin\alpha \\ &= \rho\, u_x \cos\alpha + \rho\, u_y \sin\alpha. \qquad (13)\end{aligned}$$

La composante ρu_x se transforme donc comme la variable x et l'on pourrait vérifier de même que ρu_y se transforme comme y.

On procéderait de même pour étudier les transformations de la densité et du courant dans les cas 2° et 3° (rotation autour de oy et transformation simple de Lorentz) et l'on aboutirait à la conclusion que les 4 quantités (4) sont bien les composantes d'un vecteur d'espace-temps.

On sait que si un vecteur $\vec{a}$ d'espace-temps a pour composantes d'espace a_1, a_2, a_3 et pour composante de temps a, la quantité : $c^2 a_4^2 - a_1^2 - a_2^2 - a_3^2$, qui mesure sa longueur dans l'espace-temps est un invariant indépendant du système de référence galiléen choisi. Si l'on calcule la longueur du vecteur densité courant, on trouve après un calcul sans difficultés mais un peu long :

$$c^2 \rho^2 - (\rho u_x)^2 - (\rho u_y)^2 - (\rho u_z)^2 = c^2\,|\Omega_1^2 + \Omega_2^2| \qquad (14)$$

avec :

$$\Omega_1 = \Psi_1^* \Psi_1 + \Psi_2^* \Psi_2 - \Psi_3^* \Psi_3 - \Psi_4^* \Psi_4 = \sum_1^4{}_i \Psi_i^* \,.\, \alpha_4 \Psi_i$$

$$\begin{aligned}\Omega_2 &= +i\,\Psi_1^* \Psi_3 + i\,\Psi_2^* \Psi_4 - i\,\Psi_3^* \Psi_1 - i\,\Psi_4^* \Psi_3 \\ &= \sum_1^4{}_i \Psi_i^* \,.\, \alpha_1 \alpha_2 \alpha_3 \alpha_4 \Psi_i. \qquad (15)\end{aligned}$$

Les quantités Ω_1 et Ω_2 sont des invariants de sorte que la longueur (14) du vecteur densité-courant dans l'espace-temps est bien invariante. La matrice $\alpha_1 \alpha_2 \alpha_3 \alpha_4$ utilisée dans l'expression condensée de Ω_2 se calcule aisément à partir des α_i ; elle est hermitique et a pour expression :

$$\alpha_1 \alpha_2 \alpha_3 \alpha_4 = \begin{vmatrix} 0 & 0 & +i & 0 \\ 0 & 0 & 0 & -i \\ +i & 0 & 0 & 0 \\ 0 & +i & 0 & 0 \end{vmatrix} \qquad (16)$$

3. *L'onde plane dans l'absence de champs.* — Il est très instructif de traiter le cas de l'absence de champs ($V = \vec{A} = 0$). Dans ce cas les équations de Dirac prennent la forme :

$$\begin{aligned}
&\left(\frac{h}{2\pi i}\frac{1}{c}\frac{\partial}{\partial t} + m_0 c\right)\Psi_1 - \left(\frac{h}{2\pi i}\frac{\partial}{\partial x} + i\frac{h}{2\pi i}\frac{\partial}{\partial y}\right)\Psi_4 - \frac{h}{2\pi i}\frac{\partial \Psi_3}{\partial z} = 0\\
&\left(\frac{h}{2\pi i}\frac{1}{c}\frac{\partial}{\partial t} + m_0 c\right)\Psi_2 - \left(\frac{h}{2\pi i}\frac{\partial}{\partial x} - i\frac{h}{2\pi i}\frac{\partial}{\partial y}\right)\Psi_3 + \frac{h}{2\pi i}\frac{\partial \Psi_4}{\partial z} = 0\\
&\left(\frac{h}{2\pi i}\frac{1}{c}\frac{\partial}{\partial t} - m_0 c\right)\Psi_3 - \left(\frac{h}{2\pi i}\frac{\partial}{\partial x} + i\frac{h}{2\pi i}\frac{\partial}{\partial y}\right)\Psi_2 - \frac{h}{2\pi i}\frac{\partial \Psi_1}{\partial z} = 0\\
&\left(\frac{h}{2\pi i}\frac{1}{c}\frac{\partial}{\partial t} - m_0 c\right)\Psi_4 - \left(\frac{h}{2\pi i}\frac{\partial}{\partial x} - i\frac{h}{2\pi i}\frac{\partial}{\partial y}\right)\Psi_1 + \frac{h}{2\pi i}\frac{\partial \Psi_2}{\partial z} = 0
\end{aligned} \tag{17}$$

Voyons si les équations (17) admettent pour solution l'onde plane monochromatique définie par :

$$\Psi_i = a_i\, e^{\frac{2\pi i}{h}(Wt - p_x x - p_y y - p_z z)}. \tag{18}$$

En substituant dans (17), on trouve :

$$\begin{aligned}
&\left(\frac{W}{c} + m_0 c\right) a_1 + (p_x + i\,p_y)\,a_4 + p_z\,a_3 = 0\\
&\left(\frac{W}{c} + m_0 c\right) a_2 + (p_x - i\,p_y)\,a_3 - p_z\,a_4 = 0\\
&\left(\frac{W}{c} - m_0 c\right) a_3 + (p_x + i\,p_y)\,a_2 + p_z\,a_1 = 0\\
&\left(\frac{W}{c} - m_0 c\right) a_4 + (p_x - i\,p_y)\,a_1 - p_z\,a_2 = 0
\end{aligned} \tag{19}$$

Ces équations linéaires et homogènes ne peuvent être satisfaites simultanément par des a_i non nuls que si le déterminant :

$$\begin{vmatrix}
\frac{W}{c} + m_0 c & o & p_z & p_x + i\,p_y\\
o & \frac{W}{c} + m_0 c & p_x - i\,p_y & -p_z\\
p_z & p_x + i\,p_y & \frac{W}{c} - m_0 c & o\\
p_x - i\,p_y & -p_z & o & \frac{W}{c} - m_0 c
\end{vmatrix} \tag{20}$$

est nul. Le calcul un peu long de ce déterminant montre qu'il est égal à

$$\left(\frac{W^2}{c^2} - m_0^2 c^2 - p_x^2 - p_y^2 - p_z^2\right)^2.$$

Il est nul si W, p_x, p_y, p_z sont reliés par la relation bien connue de la Mécanique relativiste :

$$\frac{W^2}{c^2} - p_x^2 - p_y^2 - p_z^2 = m_0^2 c^2 \qquad (21)$$

Supposons cette relation satisfaite [1] ; alors non seulement le déterminant (20) est nul, mais tous ses mineurs du premier ordre le sont aussi. Nous pouvons donc nous donner arbitrairement deux des quatre a_i. Donnons-nous par exemple des valeurs arbitraires A et B pour a_3 et a_4 respectivement. Les équations (19) déterminent alors a_1 et a_2 et l'on trouve :

$$a_1 = -\frac{p_z A + (p_x + i p_y) B}{\frac{W}{c} + m_0 c}; \quad a_2 = -\frac{(p_x - i p_y) A - p_z B}{\frac{W}{c} + m_0 c} \qquad (22)$$

On voit donc que l'onde plane (18) est entièrement définie si l'on connaît les amplitudes A et B.

On peut faire une remarque intéressante sur les formules (22). Dans la nouvelle Mécanique comme dans la Mécanique relativiste ancienne, nous pouvons dire que l'approximation Newtonienne est valable quand l'énergie W est peu supérieure à l'énergie au repos m_0c^2 (vitesses faibles par rapport à la vitesse de la lumière).

Il résulte de (21) que si l'approximation Newtonienne est

(1) Nous admettons pour l'instant que W est positif, c'est-à-dire que l'on a d'après (21) $\frac{W}{c} = +\sqrt{m_0^2c^2 + p_x^2 + p_y^2 + p_z^2}$. Nous aurons plus tard (chap. XX) à nous demander ce que peut signifier la solution négative $\frac{W}{c} = -\sqrt{m_0^2 c^2 + p_x^2 + p_y^2 + p_z^2}$.

valable, chacune des quantités p_x, p_y, p_z est très petite devant $m_0 c$. Reportons-nous alors aux formules (22): quand l'approximation Newtonienne est valable, leurs dénominateurs sont sensiblement égaux à $2 m_0 c$ et l'on voit que a_1 et a_2 sont très petits devant A et B. Il en résulte qu'à l'approximation Newtonienne, les fonctions Ψ_1 et Ψ_2 sont presque négligeables devant Ψ_3 et Ψ_4. On est donc ramené dans ce cas à un problème à deux Ψ_i comme dans la théorie non relativiste de Pauli. Un cas particulièrement simple où l'approximation Newtonienne se trouve rigoureusement valable est celui de l'électron au repos dans le système de référence employé ; les fonctions Ψ_i sont alors :

$$\Psi_1 = \Psi_2 = 0\ ; \quad \Psi_3 = \mathrm{A}\, e^{\frac{2\pi i}{h} m_0 c^2 t}\ ; \quad \Psi_4 = \mathrm{B}\, e^{\frac{2\pi i}{h} m_0 c^2 t} \tag{23}$$

Donc dans un système d'axes galiléens attaché à l'électron les ondes Ψ_1 et Ψ_2 sont rigoureusement nulles.

4. ***Densité et courant dans l'onde plane.*** — Nous venons de voir qu'en l'absence de champ, les équations de Dirac admettent la solution :

$$\begin{aligned}
\Psi_1 &= -\frac{p_z \mathrm{A} + (p_x + i p_y)\mathrm{B}}{\frac{\mathrm{W}}{c} + m_0 c}\, e^{\frac{2\pi i}{h}(\mathrm{W}t - p_x x - p_y y - p_z z)} \\
\Psi_2 &= -\frac{(p_x - i p_y)\mathrm{A} - p_z \mathrm{B}}{\frac{\mathrm{W}}{c} + m_0 c}\, e^{\frac{2\pi i}{h}(\mathrm{W}t - p_x x - p_y y - p_z z)} \\
\Psi_3 &= \mathrm{A}\, e^{\frac{2\pi i}{h}(\mathrm{W}t - p_x x - p_y y - p_z z)} \\
\Psi_4 &= \mathrm{B}\, e^{\frac{2\pi i}{h}(\mathrm{W}t - p_x x - p_y y - p_z z)}
\end{aligned} \tag{24}$$

les constantes W, p_x, p_y, p_z étant reliées par la relation (21). Les constantes A et B sont arbitraires sous réserve de la condition de normalisation.

Calculons la densité ρ de l'onde plane (24)

$$\rho = \sum_{1}^{4}{}_i \Psi_i^* \Psi_i = AA^* + BB^* + (AA^* + BB^*) \frac{p_x^2 + p_y^2 + p_z^2}{\left(\frac{W}{c} + m_0 c\right)^2} \tag{25}$$

$$= (AA^* + BB^*) \left(1 + \frac{\frac{W^2}{c^2} - m_0^2 c^2}{\left(\frac{W}{c} + m_0 c\right)^2} \right) = (AA^* + BB^*) \frac{2 \frac{W}{c}}{\frac{W}{c} + m_0 c}$$

Calculons de même ρu_x d'après les définitions (4)

$$\rho u_x = - c (\Psi_1^* \Psi_4 + \Psi_2^* \Psi_3 + \Psi_3^* \Psi_2 + \Psi_4^* \Psi_1)$$
$$= (AA^* + BB^*) \frac{2 p_x c}{\frac{W}{c} + m_0 c} \tag{26}$$

ce qui peut s'écrire en tenant compte de (25) :

$$\rho u_x = \rho \cdot \frac{p_x c^2}{W} \tag{27}$$

La composante u_x de la vitesse de la probabilité de présence est donc :

$$u_x = \frac{p_x c^2}{W} \tag{28}$$

Or la composante v_x de la vitesse de l'électron dans la conception classique est telle que :

$$p_x = \frac{m_0 v_x}{\sqrt{1 - \beta^2}} = \frac{W}{c^2} v_x \tag{29}$$

d'où résulte par comparaison avec (28) :

$$u_x = v_x \tag{30}$$

En raisonnant sur les composantes y et z, on obtiendrait des formules analogues à (26), (27) et (28) et on en conclurait :

$$u_y = v_y \qquad u_z = v_z \tag{31}$$

Bref la vitesse $\vec{u}$ de la probabilité dans l'onde plane est partout égale à la vitesse $\vec{v}$ que l'ancienne conception corpusculaire attribuerait à l'électron associé à cette onde plane.

Naturellement, dans la théorie de Dirac comme dans la Mécanique ondulatoire à une seule fonction Ψ, l'onde plane correspond au cas où l'on connaît exactement l'état dynamique du corpuscule ($p_x p_y p_z$ et par suite W connus) mais où sa position est totalement inconnue.

Les expressions de la densité d'électricité et de la densité de courant électrique sont ici :

$$\delta = -e\rho \; ; \; j_x = -e\rho v_x \; ; \; j_y = -e\rho v_y \; ; \; j_z = -e\rho v_z \quad (32)$$

ρ ayant la valeur (25).

CHAPITRE XIII

Le magnétisme propre de l'électron

1. *Le « globule de probabilité » en Mécanique ondulatoire.*

L'IDÉE primitive de MM. Uhlenbeck et Goudsmit consistait à considérer l'électron comme une petite sphère d'électricité en rotation autour de l'un de ses diamètres et possédant par suite (du moins dans son système propre), un moment magnétique dirigé suivant ce diamètre. Cette conception ne peut pas être conservée à la lettre dans la Mécanique nouvelle à cause de l'impossibilité où l'on se trouve d'attribuer une position et une structure à un électron. Néanmoins, nous allons voir qu'il est possible à l'aide de la fiction du « fluide de probabilité » d'obtenir dans la théorie de Dirac une sorte d'image moyenne de l'électron se rapprochant de l'image d'Uhlenbeck et Goudsmit. Pour cela, il nous faut d'abord rappeler quelques points de la Mécanique ondulatoire à une fonction Ψ.

Quand nous envisageons un mouvement d'électron qui s'opère à grande échelle, par exemple la déviation d'un électron par un champ magnétique, il nous suffit pour pouvoir décrire ce mouvement à la façon classique, de pouvoir attribuer à l'électron une localisation qui soit compatible avec les relations d'incertitude. Or il est facile de voir sur les formules

que, dans les conditions usuelles de l'expérience, la longueur d'onde de l'onde associée à l'électron est beaucoup plus petite que la plus petite longueur que nous puissions directement mesurer. Il en résulte qu'il est possible de construire un petit groupe d'ondes, formé par la superposition d'ondes planes monochromatiques de fréquences extrêmement voisines, dont les dimensions sont négligeables à notre échelle. Une observation précise faite sur un électron peut donc nous permettre, sans violer les relations d'incertitude, d'attribuer à l'électron un état de mouvement et une position pratiquement bien déterminés *à notre échelle*. Le fluide fictif de probabilité dont la densité est par définition égale à l'intensité $\Psi\Psi^*$ forme dans ce cas une sorte de petit globule à l'intérieur duquel la force due au champ appliqué peut être considérée comme constante. Il résulte alors du théorème d'Ehrenfest, puisque le globule se confond sensiblement avec son centre de gravité, que ce globule se meut comme un point matériel obéissant aux lois de la Mécanique classique. Comme le corpuscule ne peut manifester sa présence qu'à l'intérieur du globule et que les dimensions du globule sont pratiquement négligeables, tout se passe comme si le corpuscule lui-même obéissait aux lois classiques. Telle est la façon dont s'opère dans le domaine macroscopique la jonction entre l'ancienne et la nouvelle Mécanique.

Mais il faut bien remarquer que le globule de probabilité ne représente pas la structure interne de l'électron comme on avait pu le penser tout d'abord. La densité électrique $-e\rho$ à l'intérieur du globule ne représente pas une densité électrique vraie qui existerait à l'intérieur de l'électron supposé étendu. Dans les théories actuelles, on suppose l'électron ponctuel et la densité $-e\rho$ n'est, comme nous l'avons déjà expliqué, qu'une densité électrique moyenne. Le globule de probabilité n'est donc qu'une sorte d'image *moyenne* des localisations possibles de l'électron. C'est cette représentation moyenne qui se rapproche le plus dans la nouvelle Mécanique de la conception classique de l'électron. Aussi est-ce en étudiant le globule de probabilité dans la théorie de

Dirac que nous devons nous attendre à voir apparaître le magnétisme propre de l'électron sous une forme analogue à l'idée primitive de MM. Uhlenbeck et Goudsmit. Nous allons voir bientôt qu'il en est bien ainsi.

Nous allons rappeler comment on peut obtenir un modèle simple de globule de probabilité dans le cadre de la Mécanique ondulatoire à une fonction Ψ (Darwin). Nous supposerons que l'onde Ψ a, à l'instant initial $t=0$, la forme :

$$\Psi(x, y, z, 0) = a\, e^{-\frac{x^2+y^2+z^2}{2\sigma^2}}\, e^{-\frac{2\pi i}{h} m_0 (v_x x - v_y y + v_z z)} \qquad (1)$$

dont l'amplitude à symétrie sphérique autour de l'origine des coordonnées est une fonction de Gauss du rayon vecteur [1]. L'amplitude devient négligeable dès que la distance à l'origine devient un petit multiple de σ : on peut dire que le globule à l'instant initial a des dimensions de l'ordre de σ. Le second facteur exponentiel dans (1) représente le facteur de phase d'une onde plane à l'instant zéro. Pour que l'onde Ψ soit équivalente à un groupe d'ondes, la quantité σ qui mesure ses dimensions doit être petite par rapport à la longueur d'onde $\frac{h}{m_0 v}$; en raison de l'extrême petitesse de cette dernière quantité, σ peut néanmoins être négligeable à notre échelle.

M. Darwin a étudié la propagation d'un globule de la forme (1). Il a montré [2] que pendant un intervalle de temps suffisamment court le globule se transporte en bloc avec la vitesse $\vec{v}$ de telle sorte que l'on a au temps t :

$$\Psi(x, y, z, t) = a\, e^{-\frac{(x - v_x t)^2 + (y - v_y t)^2 + (z - v_z t)^2}{2\sigma^2}}$$
$$e^{\frac{2\pi i}{h}[Wt - m_0(v_x x + v_y y + v_z z)]} \qquad (2)$$

Ceci est en accord avec le théorème d'Ehrenfest. Mais avec

---

(1) Pour que Ψ soit normée, on doit avoir $|a| = \pi^{-\frac{3}{4}}\sigma^{-\frac{3}{2}}$.

(2) On trouvera le développement des calculs dans le livre de l'auteur : *Introduction à l'étude de la Mécanique ondulatoire*, Paris, Hermann, 1930, chapitre XIII.

le temps le globule a toujours une tendance à s'étaler en progressant.

Sans insister sur ce dernier point, nous pouvons dire que le globule sphérique de Darwin nous fournit une sorte d'image du mouvement macroscopique de l'électron. Il est intéressant de calculer la densité et le courant de probabilité qui lui correspondent. Pour cela nous devons nous servir des formules de la Mécanique ondulatoire à un Ψ :

$$\rho = \Psi \Psi^* \qquad \rho \vec{u} = \frac{h}{4\pi i m_0}\left[\Psi \operatorname{grad} \Psi^* - \Psi^* \operatorname{grad} \Psi\right] \tag{3}$$

On trouve aisément à l'instant initial :

$$\rho = a^2\, e^{-\frac{x^2+y^2+z^2}{\sigma^2}} \qquad \rho \vec{u} = \Psi \Psi^* \vec{v} = \rho \vec{v} \tag{4}$$

d'où

$$\vec{u} = \vec{v}$$

Le mouvement de la probabilité (ou de la répartition électrique moyenne) est donc une convection avec la vitesse $\vec{v}$ et le globule de Darwin donne bien à ce point de vue une image macroscopique moyenne de l'électron classique. En transposant le globule de Darwin dans la théorie de Dirac, nous allons voir apparaître l'électron magnétique et tournant d'Uhlenbeck et Goudsmit.

2. ***Le globule sphérique de probabilité dans la théorie de Dirac.*** — Passons maintenant à la Mécanique ondulatoire de Dirac avec 4 fonctions Ψ. Nous avons vu qu'on pouvait se donner arbitrairement les amplitudes initiales de Ψ_3 et Ψ_4 et qu'alors Ψ_1 et Ψ_2 se déduisaient des équations de propagation. Donnons-nous donc ici les Ψ_3 et Ψ_4 initiaux sous la forme inspirée de (1) :

$$\begin{aligned} \Psi_3(x, y, z, o) &= A\, e^{-\frac{x^2+y^2+z^2}{2\sigma^2}}\, e^{-\frac{2\pi i}{h} m_0 (v_x x + v_y y + v_z z)} \\ \Psi_4(x, y, z, o) &= B\, e^{-\frac{x^2+y^2+z^2}{2\sigma^2}}\, e^{-\frac{2\pi i}{h} m_0 (v_x x + v_y y + v_z z)} \end{aligned} \tag{6}$$

Supposons valable l'approximation Newtonienne : nous avons vu au dernier chapitre ce que cela signifie. Nous pou-

vons alors remplacer $\frac{h}{2\pi i}\frac{\partial \Psi'_3}{\partial t}$ et $\frac{h}{2\pi i}\frac{\partial \Psi'_4}{\partial t}$ par $m_0 c^2$ et les deux premières équations de Dirac donnent :

$$\begin{aligned} \Psi'_1 &= \frac{1}{2 m_0 c}\frac{h}{2\pi i}\left[\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)\Psi'_4 + \frac{\partial \Psi'_3}{\partial z}\right] \\ \Psi'_2 &= \frac{1}{2 m_0 c}\frac{h}{2\pi i}\left[\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)\Psi'_3 - \frac{\partial \Psi'_4}{\partial z}\right] \end{aligned} \tag{7}$$

Posons : $$P = e^{-\frac{x^2+y^2+z^2}{2\sigma^2}} e^{-\frac{2\pi i}{h} m_0 (v_x x + v_y y + v_z z)} \tag{8}$$

Nous trouvons en substituant dans (7) les valeurs des dérivées de Ψ'_3 et Ψ'_4 :

$$\Psi'_1(x, y, z, o) =$$
$$-\frac{1}{2 m_0 c}\left[A\left(m_0 v_z + \frac{h}{2\pi i}\frac{z}{\sigma^2}\right) + B\left(m_0(v_x + i v_y) + \frac{h}{2\pi i}\frac{x+iy}{\sigma^2}\right)\right] P$$
$$\Psi'_2(x, y, z, o) = \tag{9}$$
$$-\frac{1}{2 m_0 c}\left[A\left(m_0(v_x + i v_y) + \frac{h}{2\pi i}\frac{x-iy}{\sigma^2}\right) - B\left(m_0 v_z + \frac{h}{2\pi i}\frac{z}{\sigma^2}\right)\right] P$$

Ces formules donnent Ψ'_1 et Ψ'_2 à l'instant o pour une petite valeur de $\frac{v}{c}$.

Les fonctions Ψ'_1 et Ψ'_2 sont presque négligeables devant Ψ'_3 et Ψ'_4 comme nous pouvions le prévoir.

Si nous formons l'expression de la densité $\rho = \sum_1^4 \Psi'^*_i \Psi'_i$, les deux premiers termes de la somme Σ peuvent être négligés devant les deux derniers et il est suffisant d'écrire :

$$\rho = (AA^* + BB^*)\, e^{-\frac{x^2+y^2+z^2}{\sigma^2}} \tag{10}$$

Naturellement on doit avoir (normalisation):

$$(AA^* + BB^*)\int\int\int_{-\infty}^{+\infty} e^{-\frac{x^2+y^2+z^2}{\sigma^2}}\, dx\, dy\, dz = 0 \tag{11}$$

En multipliant (10) par $-e$, on obtient la densité électrique moyenne δ.

Il faut calculer maintenant les composantes de $\rho\overrightarrow{u}$. Ici

aucun terme n'est négligeable car, dans l'expression de ces quantités, les 4 termes sont les produits d'une fonction d'onde de faible valeur par une fonction d'onde de grande valeur. On a par exemple :

$$\rho u_x = -c(\Psi_1^* \Psi_4 + \Psi_2^* \Psi_3 + \Psi_3^* \Psi_2 + \Psi_4^* \Psi_1) \quad (12)$$

ce qui donne d'après (7) et (9) :

$$\rho u_x = \Big[AA^* \Big(v_x - \frac{h}{2 m_0 \pi} \frac{y}{\sigma^2} \Big) + BB^* \Big(v_x + \frac{h}{2 \pi m_0} \frac{y}{\sigma^2} \Big) + \frac{h}{2 \pi i m_0} \frac{z}{\sigma^2} (AB^* - BA^*) \Big] e^{-\frac{x^2+y^2+z^2}{\sigma^2}}$$

$$= \rho v_x + \frac{h}{2 \pi m_0} \Big[(BB^* - AA^*) \frac{y}{\sigma^2} - \frac{z}{\sigma^2} (iAB^* - iBA^*) \Big] e^{-\frac{x^2+y^2+z^2}{\sigma^2}}$$

$$\rho u_x = \rho v_x + \frac{h}{4 \pi m_0} \Big[(AA^* - BB^*) \frac{\partial}{\partial y} - (iA^*B - iAB^*) \frac{\partial}{\partial z} \Big] e^{-\frac{x^2+y^2+z^2}{\sigma^2}} \quad (13)$$

On trouve de même :

$$\rho u_y = -c(-i\Psi_1^* \Psi_4 + i\Psi_2^* \Psi_3 - i\Psi_3^* \Psi_2 + i\Psi_4^* \Psi_1)$$

$$= \rho v_y + \frac{h}{4 \pi m_0} \Big[(AB^* + BA^*) \frac{\partial}{\partial z} - (AA^* - BB^*) \frac{\partial}{\partial x} \Big] e^{-\frac{x^2+y^2+z^2}{\sigma^2}} \quad (13^{bis})$$

$$\rho u_z = -c(\Psi_1^* \Psi_3 - \Psi_2^* \Psi_1 + \Psi_3^* \Psi_1 - \Psi_4^* \Psi_1)$$

$$= \rho v_z + \frac{h}{4 \pi m_0} \Big[(iA^*B - iAB^*) \frac{\partial}{\partial x} - (AB^* + A^*B) \frac{\partial}{\partial y} \Big] e^{-\frac{x^2+y^2+z^2}{\sigma^2}}$$

Pour obtenir les composantes de la densité de courant électrique moyen, il suffit de multiplier ρu_x, ρu_y, ρu_z par $-e$ ou plus exactement, pour obtenir l'expression en u.e.m, par $-\frac{e}{c}$. Nous trouvons donc :

$$j_x = -\frac{e}{c} \rho v_x + \frac{eh}{4 \pi m_0 c} \Big[(BB^* - AA^*) \frac{\partial}{\partial y} - (iAB^* - iA^*B) \frac{\partial}{\partial z} \Big] e^{-\frac{x^2+y^2+z^2}{\sigma^2}}$$

$$j_y = -\frac{e}{c}\rho v_y + \frac{eh}{4\pi m_0 c}\left[-(A^*B + AB^*)\frac{\partial}{\partial z} - (BB^* - A^*A)\frac{\partial}{\partial x}\right] e^{-\frac{x^2+y^2+z^2}{\sigma^2}} \qquad (14)$$

$$j_z = -\frac{e}{c}\rho v_z + \frac{eh}{4\pi m_0 c}\left[(iAB^* - iA^*B)\frac{\partial}{\partial x} - (-A^*B - AB^*)\frac{\partial}{\partial y}\right] e^{-\frac{x^2+y^2+z^2}{\sigma^2}}$$

Ce sont ces formules qu'il s'agit maintenant d'interpréter.

3. ***Démonstration d'une formule d'électromagnétisme.*** — Pour parvenir à interpréter les formules (14), nous allons avoir besoin d'une formule d'électromagnétisme dont nous allons donner la démonstration dans ce paragraphe.

On sait que l'action magnétique d'un courant permanent de densité $\overrightarrow{j}$ est définie par le potentiel-vecteur :

$$\overrightarrow{A} = \int\int\int \frac{\overrightarrow{j}\, d\tau}{r} \qquad (d\tau \text{ élément de volume}) \qquad (15)$$

le champ magnétique se déduisant de $\overrightarrow{A}$ par la formule :

$$\overrightarrow{H} = \text{rot}\, \overrightarrow{A} \qquad (16)$$

Si l'on considère le courant électrique comme constitué par un faisceau d'électrons en mouvement et si δ désigne la densité moyenne de charge dans ce faisceau d'électrons, on doit poser $\overrightarrow{j} = \delta\overrightarrow{v}$, $\overrightarrow{v}$ étant la vitesse supposée uniforme des électrons. Mais si les électrons sont porteurs d'un moment magnétique, l'expression de $\overrightarrow{j}$ est à compléter en tenant compte de l'intensité d'aimantation $\overrightarrow{I}$ qui existe alors dans le faisceau. Nous voulons chercher comment $\overrightarrow{j}$ dépend de $\overrightarrow{I}$.

Soit un petit aimant formé de deux masses magnétiques $+\mu$ et $-\mu$ situées à une distance l l'une de l'autre. Si α, β, γ

sont les cosinus directeurs de l'axe du petit aimant, le moment magnétique $\vec{m}$ de l'aimant a pour composantes :

$$m_x = \alpha \mu l = \alpha m \qquad m_y = \beta m \qquad m_z = \gamma m \tag{17}$$

Le potentiel magnétique χ créé par l'aimant en un point M situé à une distance r de son centre est :

$$\chi = \frac{\mu}{r_1} - \frac{\mu}{r_2} = \mu\left(\frac{\partial \frac{1}{r}}{\partial x}\alpha l + \frac{\partial \frac{1}{r}}{\partial y}\beta l + \frac{\partial \frac{1}{r}}{\partial z}\gamma l\right) = \left(\vec{m} \, . \, \overrightarrow{\mathrm{grad}} \frac{1}{r}\right) \tag{18}$$

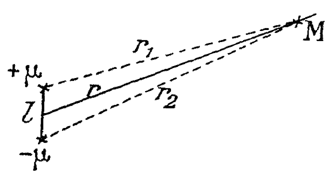


Fig. 6

et le champ créé en M a pour composantes :

$$\begin{aligned}
H_x &= -\frac{\partial \chi}{\partial x} = -\left(m_x \frac{\partial^2 \frac{1}{r}}{\partial x^2} + m_y \frac{\partial^2 \frac{1}{r}}{\partial x \, \partial y} + m_z \frac{\partial^2 \frac{1}{r}}{\partial x \, \partial z}\right) \\
H_y &= -\frac{\partial \chi}{\partial y} = -\left(m_x \frac{\partial^2 \frac{1}{r}}{\partial y \, \partial x} + m_y \frac{\partial^2 \frac{1}{r}}{\partial y^2} + m_z \frac{\partial^2 \frac{1}{r}}{\partial y \, \partial z}\right) \\
H_z &= -\frac{\partial \chi}{\partial z} = -\left(m_x \frac{\partial^2 \frac{1}{r}}{\partial z \, \partial x} + m_y \frac{\partial^2 \frac{1}{r}}{\partial z \, \partial y} + m_z \frac{\partial^2 \frac{1}{r}}{\partial z^2}\right)
\end{aligned} \tag{19}$$

Nous allons montrer que le potentiel vecteur $\vec{A}$, dont le champ H défini par (19) doit être le rotationnel, est donné par :

$$\vec{A} = \left[\vec{m} \, . \, \overrightarrow{\mathrm{grad}} \frac{1}{r}\right] \tag{20}$$

les crochets indiquant un produit extérieur. En effet, on a pour H_x :

$$H_x = (\text{rot } A)_x = \frac{\partial A_z}{\partial y} - \frac{\partial A_y}{\partial z} = \frac{\partial}{\partial y}\left(m_x \frac{\partial \frac{1}{r}}{\partial y} - m_y \frac{\partial \frac{1}{r}}{\partial x}\right)$$

$$- \frac{\partial}{\partial z}\left(m_z \frac{\partial \frac{1}{r}}{\partial x} - m_x \frac{\partial \frac{1}{r}}{\partial z}\right) = - m_y \frac{\partial^2 \frac{1}{r}}{\partial y \partial x} - m_z \frac{\partial^2 \frac{1}{r}}{\partial z \partial x}$$

$$+ m_x \left[\frac{\partial^2 \frac{1}{r}}{\partial y^2} + \frac{\partial^2 \frac{1}{r}}{\partial z^2}\right] \qquad (21)$$

d'où, en vertu de la relation bien connue $\Delta \frac{1}{r} = o$, on tire pour H_x la première expression (19). On trouverait de même les expressions (19) de H_y et H_z. C'est donc bien le potentiel vecteur (20) qui correspond au champ magnétique (19).

Supposons maintenant que nous ayons affaire non pas à un seul petit aimant de moment $\vec{m}$, mais à un corps aimanté étendu dont nous connaissons en tout point l'intensité d'aimantation $\vec{I}$. Nous avons à remplacer dans les dernières formules $\vec{m}$ par $\vec{I} d\tau$ et à intégrer, ce qui donne :

$$\vec{A} = \int\int\int \left[\vec{I} \cdot \overrightarrow{\text{grad}} \frac{1}{r}\right] d\tau \qquad (22)$$

Supposons le vecteur $\vec{I}$ nul aux limites du corps aimanté. Une intégration par parties nous permet alors d'écrire :

$$A_x = \int\int\int \left(I_y \frac{\partial \frac{1}{r}}{\partial z} - I_z \frac{\partial \frac{1}{r}}{\partial y}\right) d\tau$$

$$= \int\int_V\int \frac{1}{r}\left(\frac{\partial I_z}{\partial y} - \frac{\partial I_y}{\partial z}\right) d\tau = \int\int_V\int \frac{(\text{rot } I)_x}{r} d\tau \qquad (23)$$

(où V est le volume du corps aimanté) et nous avons des formules analogues en A_y et A_z. Nous avons donc la relation vectorielle :

$$\vec{A} = \int\int_V\int \frac{\overrightarrow{\text{rot } I}}{r} d\tau \qquad (24)$$

Si enfin nous avons affaire à un corps à la fois électrisé et aimanté en mouvement uniforme avec la vitesse $\vec{v}$, le potentiel vecteur créé par ce corps sera en vertu de (15) et de (24) :

$$\vec{A} = \iiint_V \frac{\delta \vec{v} + \vec{\mathrm{rot}\, I}}{r} d\tau \qquad (25)$$

Il est donc le même que celui qui serait créé par un courant de densité :

$$\vec{j} = \delta \vec{v} + \vec{\mathrm{rot}\, I} \qquad (26)$$

C'est cette formule (26) qui va nous servir à interpréter les formules (14).

4. *Interprétation des formules* (14). — Revenons aux formules (14) et définissons un vecteur $\vec{I}$ en nous donnant les valeurs suivantes pour ses composantes :

$$\begin{aligned} I_x &= \frac{eh}{4\pi m_0 c}(-A^*B - AB^*)\, e^{-\frac{x^2+y^2+z^2}{\sigma^2}} \\ I_y &= \frac{eh}{4\pi m_0 c}(i\, AB^* - i\, A^*B)\, e^{-\frac{x^2+y^2+z^2}{\sigma^2}} \qquad (27) \\ I_z &= \frac{eh}{4\pi m_0 c}(BB^* - AA^*)\, e^{-\frac{x^2+y^2+z^2}{\sigma^2}} \end{aligned}$$

Le vecteur $\vec{I}$ est réel (car $I_x = I_x^*$ etc.) et il est nul aux limites du globule de probabilité, c'est-à-dire ici à l'infini. Avec ce vecteur $\vec{I}$, les formules (14) peuvent s'écrire vectoriellement :

$$\vec{j} = -\frac{e}{c}\rho \vec{v} + \vec{\mathrm{rot}\, I} \qquad (28)$$

Cette formule (28) est d'un haut intérêt car on voit en la comparant avec la formule (26) que le globule de probabilité, représentation macroscopique moyenne de l'électron n'est pas assimilable à une simple sphère de charge $-e$ en

mouvement avec la vitesse $\vec{v}$, mais à une sphère à la fois électrisée et aimantée, d'intensité d'aimantation égale à $\vec{I}$ en chacun de ses points.

Le moment magnétique total du globule est le vecteur $\vec{\mathfrak{M}}$ que l'on obtient en intégrant le vecteur $\vec{I}$. Compte tenu de (27) et de la condition de normalisation (11), on trouve pour les composantes de $\vec{\mathfrak{M}}$ (¹) :

$$\begin{aligned}\mathfrak{M}_x &= \iiint I_x\, d\tau \\ &= \frac{eh}{4\pi m_0 c}(-A^*B - AB^*)\iiint e^{-\frac{x^2+y^2+z^2}{\sigma^2}}\, d\tau \\ &= -\frac{eh}{4\pi m_0 c}\frac{A^*B + AB^*}{AA^* + BB^*} \qquad (29)\end{aligned}$$

$$\mathfrak{M}_y = \iiint I_y\, d\tau = \frac{eh}{4\pi m_0 c}\,\frac{i(AB^* - A^*B)}{AA^* + BB^*}$$

$$\mathfrak{M}_z = \iiint I_z\, d\tau = \frac{eh}{4\pi m_0 c}\cdot\frac{BB^* - AA^*}{AA^* + BB^*}.$$

La longueur du vecteur $\vec{\mathfrak{M}}$ est donc :

$$\begin{aligned}|\mathfrak{M}| &= \sqrt{\mathfrak{M}_x^2 + \mathfrak{M}_y^2 + \mathfrak{M}_z^2} \\ &= \frac{eh}{4\pi m_0 c}\sqrt{\frac{(A^*B + AB^*)^2 - (A^*B - AB^*)^2 + (BB^* - AA^*)^2}{(AA^* + BB^*)^2}} \\ &= \frac{eh}{4\pi m_0 c}. \qquad (30)\end{aligned}$$

Le globule a donc un moment magnétique égal à un magnéton de Bohr.

Reprenons maintenant la formule (13) et cherchons à l'écrire sous la forme : $\rho u_x = \rho v_x + \rho v_x'$. Nous sommes amenés à poser

(¹) En réalité, $\mathfrak{M}_x$ $\mathfrak{M}_y$ $\mathfrak{M}_z$ sont seulement des valeurs moyennes comme nous l'indiquerons d'une façon précise au chapitre suivant. On devrait donc écrire $\overline{\mathfrak{M}_x}$ au lieu de $\mathfrak{M}_x$ etc.

$$v_x' = \frac{h}{2\pi m_0 \rho}\left[(\mathrm{BB}^* - \mathrm{AA}^*)\frac{y}{\sigma^2} - (i\,\mathrm{AB}^*\; i\;\mathrm{A}^*\mathrm{B})\frac{z}{\sigma^2}\right] e^{\frac{x^2+y^2+z^2}{\sigma^2}}$$

$$= \frac{h}{2\pi m_0 \sigma^2}\left[\frac{\mathrm{BB}^* - \mathrm{AA}^*}{\mathrm{AA}^* + \mathrm{BB}^*}\, y - \frac{i\,\mathrm{AB}^* - i\,\mathrm{A}^*\mathrm{B}}{\mathrm{AA}^* + \mathrm{BB}^*}\, z\right] \qquad (31)$$

en tenant compte de (10). De même nous poserons :

$$v_y' = \frac{h}{2\pi m_0 \sigma^2}\left[\frac{-(\mathrm{AB}^* + \mathrm{A}^*\mathrm{B})}{\mathrm{AA}^* + \mathrm{BB}^*}\, z - \frac{(\mathrm{BB}^* - \mathrm{AA}^*)}{\mathrm{AA}^* + \mathrm{BB}^*}\, x\right] \qquad (31^{bis})$$

$$v_z' = \frac{h}{2\pi m_0 \sigma^2}\left[\frac{i\,\mathrm{AB}^* - i\,\mathrm{A}^*\mathrm{B}}{\mathrm{AA}^* + \mathrm{BB}^*}\, x - \frac{-(\mathrm{AB}^* + \mathrm{A}^*\mathrm{B})}{\mathrm{AA}^* + \mathrm{BB}^*}\, y\right].$$

Posons alors :

$$\omega_x = \frac{h}{2\pi m_0 \sigma^2}\,\frac{(\mathrm{AB}^* + \mathrm{A}^*\mathrm{B})}{\mathrm{AA}^* + \mathrm{BB}^*}\,; \quad \omega_y = \frac{-h}{2\pi m_0 \sigma^2}\,\frac{i\,\mathrm{AB}^* - i\,\mathrm{A}^*\mathrm{B}}{\mathrm{AA}^* + \mathrm{BB}^*}\,;$$

$$\omega_z = \frac{-h}{2\pi m_0 \sigma^2}\cdot\frac{\mathrm{BB}^* - \mathrm{AA}^*}{\mathrm{AA}^* + \mathrm{BB}^*}\,; \qquad (32)$$

les formules (31) et (31*bis*) deviennent :

$$v_x' = \omega_y z - y\,\omega_z\,;\quad v_y' = \omega_z x - \omega_x z\,;\quad v_z' = \omega_x y - \omega_y x. \qquad (33)$$

Comme d'après la manière même dont nous avons introduit $\vec{v'}$, on a : $\vec{u} = \vec{v} + \vec{v'}$, on voit que la vitesse de la probabilité est la somme de la vitesse $\vec{v}$ de translation d'ensemble du globule et d'une vitesse *interne* $\vec{v'}$ due à une rotation d'ensemble définie par la vitesse de rotation $\vec{\omega}$ de composantes $\omega_x, \omega_y, \omega_z$. Cette vitesse de rotation est représentée par le vecteur $\vec{\omega}$ passant par le centre du globule, parallèle à $\vec{\mathfrak{M}}$ (car on a $\omega_x : \omega_y : \omega_z = \mathfrak{M}_x : \mathfrak{M}_y : \mathfrak{M}_z$) et de longueur

$$|\omega| = \sqrt{\omega_x^2 + \omega_y^2 + \omega_z^2} = \frac{h}{2\pi m_0 \sigma^2} \qquad (34)$$

Cette rotation interne d'ensemble du fluide électrique moyen explique l'origine du moment magnétique $\mathfrak{M}$: elle est d'autant plus rapide que le globule est plus petit, ce qui

s'explique aisément puisque le moment magnétique $\overrightarrow{\mathcal{M}}$ doit toujours être égal à un magnéton de Bohr. Le globule de probabilité dans la théorie de Dirac donne donc une sorte d'image macroscopique moyenne de l'électron magnétique tournant.

CHAPITRE XIV

Le tenseur « densité de moment électrique et de moment magnétique »

1. Le moment magnétique de l'électron de Dirac à l'approximation Newtonienne.

DANS le dernier chapitre, en étudiant le globule sphérique de probabilité défini par les fonctions d'onde (6) et (9), nous sommes parvenus à mettre en évidence un vecteur $\vec{I}$ donné par les formules 27, vecteur qui représente l'intensité d'aimantation, c'est-à-dire la densité de moment magnétique, dans le globule. Or, si l'on tient compte des formules (6), on peut écrire les formules (27) donnant les composantes de $\vec{I}$ sous la forme :

$$\begin{aligned} I_x &= \frac{e\,h}{4\,\pi\,m_0\,c}\,(-\,\Psi_3^*\,\Psi_4 - \Psi_3\,\Psi_4^*) \\ I_y &= \frac{e\,h}{4\,\pi\,m_0\,c}\,(i\,\Psi_3\,\Psi_4^* - i\,\Psi_4\,\Psi_3^*) \qquad (1) \\ I_z &= \frac{e\,h}{4\,\pi\,m_0\,c}\,(\Psi_4^*\,\Psi_4 - \Psi_3^*\,\Psi_3). \end{aligned}$$

Sous cette nouvelle forme, les expressions des composantes du vecteur $\vec{I}$ sont valables pour toute onde de Dirac à l'approximation Newtonienne (c'est-à-dire quand Ψ_1 et Ψ_2 sont

négligeables devant Ψ_3 et Ψ_4) et non plus seulement pour le globule sphérique de probabilité envisagé au dernier chapitre.

Les composantes du moment magnétique moyen de l'électron de Dirac s'obtiennent alors en intégrant les expressions (1) dans tout l'espace. On a donc :

$$\overline{\mathfrak{M}_x} = \iiint I_x \, d\tau = \frac{eh}{4\pi m_0 c} \iiint (-\Psi_3^* \Psi_4 - \Psi_4^* \Psi_3) \, d\tau$$

$$\overline{\mathfrak{M}_y} = \iiint I_y \, d\tau = \frac{eh}{4\pi m_0 c} \iiint (-i\Psi_3^* \Psi_4 + i\Psi_4^* \Psi_3) \, d\tau$$

$$\overline{\mathfrak{M}_z} = \iiint I_z \, d\tau = \frac{eh}{4\pi m_0 c} \iiint (\Psi_4^* \Psi_4 - \Psi_3^* \Psi_3) \, d\tau. \quad (2)$$

Il ne faut pas oublier que toutes ces formules ont seulement une signification statistique. Si l'on considère un très grand nombre d'électrons se trouvant tous dans un même état défini par les mêmes fonctions Ψ_3 et Ψ_4 et si l'on mesure pour chacun de ces électrons par exemple la composante le long de l'axe x du moment magnétique propre, on trouvera suivant les cas des résultats différents, mais la valeur moyenne des résultats obtenus dans l'ensemble sera le $\overline{\mathfrak{M}_x}$ des formules (2). Ceci n'est qu'une application des idées générales de la nouvelle Mécanique ; en particulier on se souviendra qu'en vertu d'une remarque faite au chapitre VI paragraphe 3, les expressions (1) ne sont pas de véritables densités physiques au sens ancien, mais sont les quantités qu'il faut intégrer pour obtenir les valeurs moyennes (2).

La troisième formule (2) doit retenir un peu notre attention. Nous savons que si l'on mesure le moment magnétique propre de l'électron parallèlement à l'axe des z on doit trouver nécessairement : $\pm$ un magnéton de Bohr. Les équations de Dirac font jouer un rôle particulier à l'axe des z précisément parce que les probabilités de ces deux hypothèses doivent s'exprimer simplement à l'aide des Ψ_i. Si alors on jette les yeux sur la formule (2), on voit que l'hypothèse de trouver pour $\mathfrak{M}_z$ la valeur $+\dfrac{eh}{4\pi m_0 c}$ est $\iiint \Psi_4^* \Psi_4 \, d\tau$

tandis que celle de trouver $-\frac{eh}{4\pi m_0 c}$ est $\int\int\int \Psi_3^* \Psi_3 \, d\tau$.
Ceci est parfaitement d'accord avec les idées de Pauli auxquelles la théorie de Dirac se ramène quand on admet comme ici l'approximation Newtonienne. La valeur moyenne (2) de $\mathfrak{M}_z$ représente bien le résultat moyen de la mesure de $\mathfrak{M}_z$ pour un grand nombre d'électrons dans l'état défini par Ψ_3 et Ψ_4.

Les expressions des composantes de $\overrightarrow{I}$, c'est-à-dire les expressions (1), sont invariantes de forme pour tout changement de coordonnées d'espace. Cela veut dire que si l'on passe d'un système de coordonnées rectangulaires xyz à un autre système rectangulaire $x'\,y'\,z'$, les composantes de $\overrightarrow{I}$ dans le nouveau système s'expriment à l'aide des nouvelles fonctions d'onde Ψ_i' comme les anciennes composantes de $\overrightarrow{I}$ s'exprimaient par (1) à l'aide des Ψ_i. Nous ne ferons pas ici la vérification qui est facile.

Mais cette invariance de forme n'a plus lieu pour une transformation de Lorentz. Ainsi se marque le caractère d'approximation Newtonienne des expressions (1). Pour obtenir une invariance relativiste, il faudra tenir compte des fonctions Ψ_1 et Ψ_2 et faire rentrer les trois composantes de l'intensité d'aimantation $\overrightarrow{I}$ parmi les six composantes d'un tenseur antisymétrique du second ordre comme nous le verrons.

2. *Moment magnétique moyen d'une onde plane à l'approximation Newtonienne.* — Considérons une onde plane définie à l'approximation Newtonienne par les deux fonctions d'onde :

$$\begin{aligned} \Psi_3 &= A\, e^{\frac{2\pi i}{h}\left[Wt - m_0 v_x x - m_0 v_y y - m_0 v_z z\right]} \\ \Psi_4 &= B\, e^{\frac{2\pi i}{h}\left[Wt - m_0 v_x x - m_0 v_y y - m_0 v_z z\right]} \end{aligned} \qquad (3)$$

quelles sont les composantes du moment magnétique moyen ?

On trouve aisément à l'aide des formules (2), compte tenu de la condition de normalisation [1]:

$$\overline{\mathfrak{M}}_x = \frac{eh}{4\pi m_0 c} \cdot \frac{-A^*B - AB^*}{AA^* + BB^*}$$
$$\overline{\mathfrak{M}}_y = \frac{eh}{4\pi m_0 c} \cdot \frac{i\,AB^* - i\,A^*B}{AA^* + BB^*} \qquad (4)$$
$$\overline{\mathfrak{M}}_z = \frac{eh}{4\pi m_0 c} \cdot \frac{BB^* - AA^*}{AA^* + BB^*}.$$

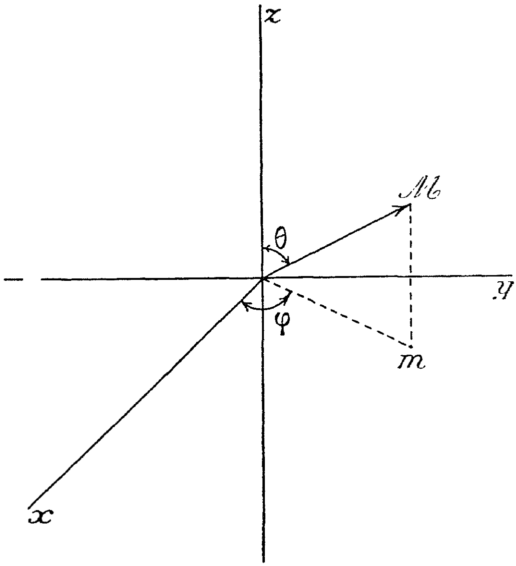


Fig. 7

[1] On peut être gêné pour écrire la condition de normalisation parce que le domaine d'intégration est infini et, rigoureusement parlant, il faudrait introduire les différentielles propres. Mais on peut pratiquement tourner la difficulté en supposant d'abord le domaine fini et de volume V. La condition de normalisation s'écrira alors ici :

$$\int\int\int (|\Psi_3|^2 + |\Psi_4|^2)\, d\tau = (AA^* + BB^*)V = 1$$

et de (2), ou tirera par exemple :

$$\overline{\mathfrak{M}}_x = \frac{eh}{4\pi m_0 c}\left(-A^*B - B^*A\right)V = \frac{eh}{4\pi m_0 c}\,\frac{-A^*B - B^*A}{AA^* + BB^*}$$

Comme ce résultat est exact, quelque grand que soit V, les formules (4) se trouvent justifiées.

Les formules (4) coïncident d'ailleurs avec les formules (29) du dernier chapitre. Nous pouvions nous y attendre car lesdites formules (29) sont valables quelle que soit la valeur de σ dans les formules (6) et si nous faisons tendre σ vers l'infini, le globule sphérique du dernier chapitre tend vers l'onde plane (3).

Repérons la direction du vecteur $\overrightarrow{\mathfrak{M}}$ dans un système de coordonnées sphériques θ et φ.

Nous avons alors :

$$\begin{aligned}\overline{\mathfrak{M}_x} : \overline{\mathfrak{M}_y} : \overline{\mathfrak{M}_z} &= (-A^*B - AB^*) : (i\,AB^* - i\,A^*\,B) \\ &\qquad : (BB^* - AA^*) \qquad (5) \\ &= \sin\theta\cos\varphi : \sin\theta\sin\varphi : \cos\theta\end{aligned}$$

d'où l'on tire aisément :

$$\left(-\frac{B}{A} - \frac{B^*}{A^*}\right) : i\left(-\frac{B}{A} + \frac{B^*}{A^*}\right) : \left(\frac{B}{A} \cdot \frac{B^*}{A^*} - 1\right) = \\ = 2\,\mathrm{ctg}\,\frac{\theta}{2}\cos\varphi : 2\,\mathrm{ctg}\,\frac{\theta}{2}\sin\varphi : \left(\mathrm{ctg}^2\frac{\theta}{2} - 1\right). \qquad (6)$$

Or on satisfait aux équations (6) en posant :

$$-\frac{B}{A} = \mathrm{ctg}\,\frac{\theta}{2}\,e^{i\varphi} \qquad (7)$$

et la relation complexe (7) équivalente à deux relations réelles nous apprend comment l'orientation de $\overrightarrow{\mathfrak{M}}$ est reliée à la valeur du rapport $\frac{B}{A}$.

On peut exprimer cette liaison de la façon suivante (Darwin, Jordan) : Considérons une sphère de rayon unité : l'orientation du vecteur $\overrightarrow{\mathfrak{M}}$ est définie par les coordonnées θ et φ du point M où ce vecteur perce la sphère. Projetons ce point M stéréographiquement sur le plan de l'équateur, le centre de projection étant au pôle nord.

Le point projection m a pour coordonnées :

$$x = \mathrm{ctg}\,\frac{\theta}{2}\cos\varphi \qquad y = \mathrm{ctg}\,\frac{\theta}{2}\sin\varphi \qquad (8)$$

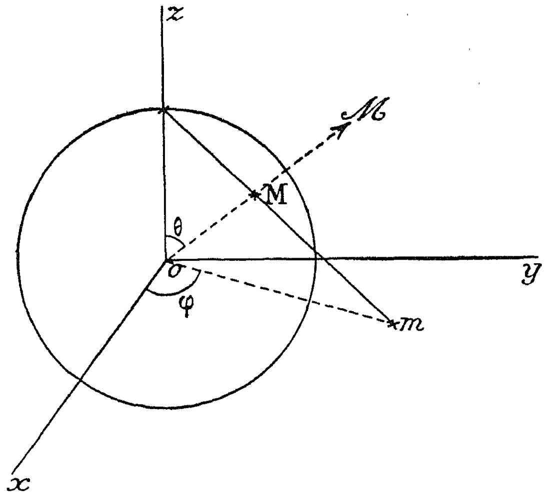


Fig. 8

Si l'on considère le plan des xy comme le plan d'une variable complexe, l'affixe du point m est $\cot g \frac{\theta}{2}\, \mathrm{e}^{i\varphi}$ et est égal à $-\frac{B}{A}$ d'après (7). D'où le résultat suivant : le rapport $-\frac{B}{A}$ est relié à la direction du moment magnétique par la même relation qui lie l'affixe d'un point dans le plan d'une variable complexe avec la direction qui lui correspond dans la sphère par projection stéréographique.

3. ***Le tenseur « Densité de moment magnétique et de moment électrique ».*** — Nous avons vu que les expressions (1) possédaient l'invariance de forme pour les changements de coordonnées d'espace, mais ne la possédaient point pour les transformations de Lorentz. Il est néanmoins possible de trouver 6 combinaisons quadratiques des 4 fonctions Ψ_i qui se transforment lors des changements de coordonnées galiléennes comme les composantes d'un tenseur antisymétrique du second ordre. Voici leurs expressions qui sont toutes réelles :

$$\mu_{yz} = \mathrm{I}_x = \frac{eh}{4\pi m_0 c}\left(\Psi'_1{}^* \Psi'_2 + \Psi'_2{}^* \Psi'_1 - \Psi'_3{}^* \Psi'_4 - \Psi'_4{}^* \Psi'_3\right) =$$

$$= \frac{eh}{4\pi m_0 c}\, i \sum_1^4{}_k \Psi'_k{}^* \cdot \alpha_2 \alpha_3 \alpha_4 \Psi'_k$$

$$\mu_{zx} = \mathrm{I}_y = \frac{eh}{4\pi m_0 c}\left(i\Psi'_1{}^* \Psi'_2 - i\Psi'_2{}^* \Psi'_1 - i\Psi'_3{}^* \Psi'_4 + i\Psi'_4{}^* \Psi'_3\right) =$$

$$= \frac{eh}{4\pi m_0 c}\, i \sum_1^4{}_k \Psi'_k{}^* \cdot \alpha_3 \alpha_1 \alpha_4 \Psi'_k$$

$$\mu_{xy} = \mathrm{I}_z = \frac{eh}{4\pi m_0 c}\left(\Psi'_1{}^* \Psi'_1 - \Psi'_2{}^* \Psi'_2 - \Psi'_3{}^* \Psi'_3 + \Psi'_4{}^* \Psi'_4\right) =$$

$$= \frac{eh}{4\pi m_0 c}\, i \sum_1^4{}_k \Psi'_k{}^* \cdot \alpha_1 \alpha_2 \alpha_4 \Psi'_k$$

$$\mu_{xt} = \mathrm{J}_x = \frac{eh}{4\pi m_0 c}\left(i\Psi'_1{}^* \Psi'_4 + i\Psi'_2{}^* \Psi'_3 - i\Psi'_3{}^* \Psi'_2 - i\Psi'_4{}^* \Psi'_1\right) =$$

$$= \frac{eh}{4\pi m_0 c}\, i \sum_1^4{}_k \Psi'_k \cdot \alpha_1 \alpha_4 \Psi'_k$$

$$\mu_{yt} = \mathrm{J}_y = \frac{eh}{4\pi m_0 c}\left(-\Psi'_1{}^* \Psi'_4 + \Psi'_2{}^* \Psi'_3 + \Psi'_3{}^* \Psi'_2 - \Psi'_4{}^* \Psi'_1\right) =$$

$$= \frac{eh}{4\pi m_0 c}\, i \sum_1^4{}_k \Psi'_k{}^* \cdot \alpha_2 \alpha_4 \Psi'_k$$

$$\mu_{zt} = \mathrm{J}_z = \frac{eh}{4\pi m_0 c}\left(i\Psi'_1{}^* \Psi'_3 - i\Psi'_2{}^* \Psi'_4 - i\Psi'_3{}^* \Psi'_1 + i\Psi'_4{}^* \Psi'_2\right) =$$

$$= \frac{eh}{4\pi m_0 c}\, i \sum_1^4{}_k \Psi'_k{}^* \cdot \alpha_3 \alpha_4 \Psi'_k \qquad (9)$$

A l'aide des 6 quantités (9) et en posant les relations d'antisymétrie $\mu_{zy} = -\mu_{yz}$ etc., on peut constituer un tableau antisymétrique à 4 rangs et 4 colonnes :

$$\begin{vmatrix} 0 & \mu_{xy} & \mu_{xz} & \mu_{xt} \\ \mu_{yx} & 0 & \mu_{yz} & \mu_{yt} \\ \mu_{zx} & \mu_{zy} & 0 & \mu_{zt} \\ \mu_{tx} & \mu_{ty} & \mu_{tz} & 0 \end{vmatrix} \qquad (10)$$

Il est facile de vérifier que le tableau (10) définit un tenseur antisymétrique du second ordre. Pour cela, il suffit de vérifier que pour chacune des trois transformations simples

en lesquelles on peut décomposer une transformation générale de Lorentz, les μ_{ij} se transforment suivant le schéma :

$$\mu_{ij}' = \sum_{kl} \mu_{kl} \frac{dx_i'}{dx_k} \frac{dx_j'}{dx_l} \tag{11}$$

Nous passons sous silence cette vérification qui est longue et sans intérêt.

Si l'on peut négliger les fonctions d'onde Ψ_1 et Ψ_2, ce qui arrive quand l'approximation Newtonienne est valable, les quantités μ_{yz}, μ_{zx} et μ_{xy} se réduisent respectivement aux I_x, I_y et I_z de la formule (1). Il est donc naturel de penser que ces trois composantes de μ sont les trois composantes de l'aimantation magnétique moyenne $\vec{I}$ dans leurs expressions relativistes exactes.

Mais que représentent les 3 composantes μ_{xt}, μ_{yt} et μ_{zt} ? En théorie de Relativité, on doit unir le champ magnétique et le champ électrique pour définir un tenseur antisymétrique du second ordre. Dans un mémoire sur l'électron magnétique antérieur à la théorie de Dirac, M. Frenkel a montré la nécessité de compléter le moment magnétique de l'électron par un moment électrique (1). Voici comment Frenkel a raisonné. Dans la théorie ancienne de l'électron magnétique, l'électron est un corpuscule possédant un moment magnétique propre ; dans un système de référence où l'électron est en mouvement, il doit créer autour de lui un champ électrique dû à son moment magnétique car, de même qu'une charge électrique en mouvement est équivalente à un courant et crée autour d'elle un champ magnétique, un pôle magnétique en mouvement crée autour de lui un champ électrique. L'électron magnétique en mouvement doit donc posséder un moment électrique et, pour satisfaire aux exigences du principe de Relativité, les trois composantes du moment électrique doivent s'unir aux trois composantes du moment magnétique pour permettre de définir un tenseur

(1) *Zeitschrift für Physik*, 37, 4-5, p. 243.

antisymétrique du second ordre. Nous sommes donc amenés à considérer μ_{xt} μ_{yt} μ_{zt} comme les trois composantes d'une densité de moment électrique.

Les 6 quantités (9) permettent de former 2 invariants c'est-à-dire qu'il existe 2 combinaisons des μ_{ij} dont la valeur est la même dans tous les systèmes de référence. Ces deux combinaisons sont les suivantes :

$$I^2 - J^2 = I_x^2 + I_y^2 + I_z^2 - J_x^2 - J_y^2 - J_z^2 ;$$
$$(\vec{I} \cdot \vec{J}) = I_x J_x + I_y J_y + I_z J_z . \tag{12}$$

En effet introduisons les deux invariants déjà rencontrés au chapitre XII, paragraphe 3 :

$$\Omega_1 = \Psi_1^* \Psi_1 + \Psi_2^* \Psi_2 - \Psi_3^* \Psi_3 - \Psi_4^* \Psi_4 = \sum_1^4{}_k \Psi_k^* \cdot \alpha_4 \Psi_k$$
$$\Omega_2 = i \Psi_1^* \Psi_3 + i \Psi_2^* \Psi_4 - i \Psi_3^* \Psi_1 - i \Psi_4^* \Psi_2 =$$
$$= \sum_1^4{}_k \Psi_k^* \cdot \alpha_1 \alpha_2 \alpha_3 \alpha_4 \Psi_k \tag{13}$$

On vérifie aisément les relations :

$$I^2 - J^2 = \left(\frac{eh}{4 \pi m_0 c}\right)^2 (\Omega_1^2 - \Omega_2^2) ; (\vec{I} \cdot \vec{J}) = \left(\frac{eh}{4 \pi m_0 c}\right)^2 \Omega_1 \Omega_2 \tag{14}$$

qui montrent bien l'invariance des deux quantités. Cette invariance est d'ailleurs évidente au point de vue tensoriel car d'une part $I^2 - J^2$ est la « longueur » du tenseur dans l'espace temps et d'autre part le produit scalaire

$$\left(\vec{I} \cdot \vec{J}\right) = \mu_{yz} \mu_{zt} + \mu_{zx} \mu_{yt} + \mu_{xy} \mu_{zt}$$

qui est évidemment un invariant des changements de coordonnées d'espace, ne varie pas non plus pour une transformation simple de Lorentz comme il est aisé de le vérifier ; d'où résulte son invariance pour une transformation générale de Lorentz.

Comme précédemment nous pouvons remarquer que les expressions (9) ne sont pas des densités physiques au sens

ancien du mot : ce sont seulement les quantités qu'il faut intégrer dans l'espace pour obtenir les valeurs moyennes des composantes du moment magnétique et du moment électrique de l'électron. Ces deux moments *moyens* que nous désignerons par $\vec{\mathfrak{M}}$ et $\vec{\mathfrak{P}}$ sont donnés par les formules :

$$\vec{\mathfrak{M}} = \int\int\int \vec{\mathrm{I}}\, d\tau \;; \qquad \vec{\mathfrak{P}} = \int\int\int \vec{\mathrm{J}}\, d\tau \tag{15}$$

les vecteurs $\vec{\mathrm{I}}$ et $\vec{\mathrm{J}}$ étant définis par leurs composantes (9). Naturellement les valeurs des composantes $\vec{\mathfrak{M}}_x$ etc. doivent être interprétées statistiquement. Ainsi la formule :

$$\vec{\mathfrak{M}}_z = \int\int\int \mathrm{I}_z\, d\tau = \frac{eh}{4\pi m_0 c}$$

$$\int\int\int (\Psi_1^* \Psi_1 - \Psi_2^* \Psi_2 - \Psi_3^* \Psi_3 + \Psi_4^* \Psi_4)\, d\tau \tag{16}$$

signifie qu'une mesure de $\mathfrak{M}_z$ peut fournir la valeur $+\frac{eh}{4\pi m_0 c}$ avec une probabilité $\int\int\int (\Psi_1^* \Psi_1 + \Psi_4^* \Psi_4)\, d\tau$ et la valeur $-\frac{eh}{4\pi m_0 c}$ avec une probabilité

$$\int\int\int (\Psi_2^* \Psi_2 + \Psi_3^* \Psi_3) d\tau.$$

La somme des deux probabilités est égale à l'unité en vertu de la condition de normalisation des Ψ_k.

M. Frenkel (1), en raisonnant sur l'ancienne image purement corpusculaire de l'électron, avait démontré que le moment magnétique $\vec{\mathfrak{M}}$ de l'électron dans un système où sa vitesse est $\vec{v}$ doit être relié à son moment électrique $\vec{\mathfrak{P}}$ dans le même système par la relation :

$$\vec{\mathfrak{P}} = \left[\vec{\mathfrak{M}} \cdot \frac{\vec{v}}{c}\right] \tag{17}$$

(1) *Loc. cit.*

Nous allons voir sur un exemple que, dans la théorie de Dirac, la relation de Frenkel reste exacte.

4. *Exemple simple : l'onde plane monochromatique.* — Servons-nous des formules du chapitre XII, paragraphe 3, en prenant, pour simplifier, la direction de propagation de l'onde plane comme axe des z. On a alors : $p_x = p_y = 0$ et les 4 Ψ'_i ont la forme :

$$\Psi'_3 = A\, e^{\frac{2\pi i}{h}(Wt - p_z z)}; \qquad \Psi'_4 = B\, e^{\frac{2\pi i}{h}(Wt - p_z z)}$$

$$\Psi'_1 = -\frac{p_z A}{\frac{W}{c} + m_0 c}\, e^{\frac{2\pi i}{h}(Wt - p_z z)}; \quad \Psi'_2 = \frac{p_z B}{\frac{W}{c} + m_0 c}\, e^{\frac{2\pi i}{h}(Wt - p_z z)} \tag{18}$$

Avec cette forme des Ψ'_i, les formules (9) donnent aisément :

$$I_x = -\frac{eh}{4\pi m_0 c}(A^*B + AB^*)\left(1 + \frac{p_z^2}{\left(\frac{W}{c} + m_0 c\right)^2}\right);$$

$$J_x = \frac{eh}{4\pi m_0 c} \cdot \frac{2 i p_z}{\frac{W}{c} + m_0 c}(AB^* - A^*B)$$

$$I_y = i\frac{eh}{4\pi m_0 c}(AB^* - BA^*)\left(1 + \frac{p_z^2}{\left(\frac{W}{c} + m_0 c\right)^2}\right);$$

$$J_y = \frac{eh}{4\pi m_0 c} \cdot \frac{2 p_z}{\frac{W}{c} + m_0 c}(A^*B + AB^*) \tag{19}$$

$$I_z = \frac{eh}{4\pi m_0 c}(BB^* - AA^*)\left(1 - \frac{p_z^2}{\left(\frac{W}{c} + m_0 c\right)^2}\right); \quad J_z = 0.$$

On vérifie sans peine que l'on a :

$$(\vec{I} \cdot \vec{J}) = 0. \tag{20}$$

Ceci est d'ailleurs une conséquence de la seconde formule (14) car l'invariant Ω_2 est nul dans le cas actuel [1].

L'examen des formules (19) fournit les relations :

$$J_x = I_y \frac{2\,p_z\left(\frac{W}{c} + m_0\,c\right)}{\left(\frac{W}{c} + m_0\,c\right)^2 + p_z^2} = I_y \frac{v_z}{c};$$

$$J_y = -\,I_x \frac{2\,p_z\left(\frac{W}{c} + m_0\,c\right)}{\left(\frac{W}{c} + m_0\,c\right)^2 + p_z^2} = -\,I_x \frac{v_z}{c}; \qquad J_z = 0. \quad (21)$$

En intégrant les relations (21) dans tout l'espace pour faire apparaître les composantes de $\vec{\mathfrak{M}}$ et de $\vec{\mathfrak{L}}$, on voit que la relation (17) de Frenkel est vérifiée (car $v_x = v_y = 0$).

Quand l'approximation Newtonienne est valable, le rapport $\dfrac{p_z}{\frac{W}{c} + m_0 c}$ est sensiblement égal à $\dfrac{m_0\,v_z}{2\,m_0\,c} = \dfrac{1}{2}\,\dfrac{v_z}{c}$ et son carré est négligeable devant l'unité. On trouve alors pour $I_x\,I_y\,I_z$ les valeurs que l'on obtient à partir des formules (27) du chapitre précédent en y faisant tendre σ vers l'infini, résultat auquel on pouvait s'attendre.

Sur les formules générales (19) et (20), on peut faire les remarques suivantes : le vecteur $\vec{J}$ est normal au plan déterminé par le vecteur $\vec{I}$ et le vecteur $\vec{p}$; si la vitesse tend vers la vitesse de la lumière, c'est-à-dire si p_z tend vers $\dfrac{W}{c}$ (qui est alors beaucoup plus grand que $m_0\,c$), le trièdre formé par les trois vecteurs $\vec{p}$, $\vec{I}$ et $\vec{J}$ tend à devenir trirectangle tandis que les longueurs de $\vec{I}$ et de $\vec{J}$ tendent à devenir égales.

---

[1] En effet, on peut évidemment calculer cet invariant dans n'importe quel système Galiléen, par exemple dans un système où la vitesse de l'électron est nulle ; or, dans un tel système $\Psi_1 = \Psi_2 = 0$ et (13) donne $\Omega_2 = 0$.

CHAPITRE XX

Matrices et intégrales premières dans la théorie de Dirac
Moment de rotation propre de l'électron

1. *Les valeurs et fonctions propres des équations de Dirac.*

LES valeurs et fonctions propres des équations de Dirac se définissent aisément par analogie avec le cas de la théorie à une fonction Ψ.

Quand le champ extérieur est indépendant du temps, il existe des solutions monochromatiques des équations de Dirac c'est-à-dire des solutions où les quatre Ψ_k ne dépendent du temps que par un même facteurs exponentiel $e^{\frac{2\pi i}{h}Wt}$. Les quatre Ψ_k satisfont alors les équations :

$$\left[\left(\frac{W}{c}+\frac{eV}{c}\right)+\alpha_1 P_1+\alpha_2 P_2+\alpha_3 P_3+\alpha_4 m_0 c\right]\Psi_k=0;\ (k=1,2,3,4). \tag{1}$$

Les valeurs de la constante W pour lesquelles il existe au moins un ensemble de Ψ_k qui soient finies, uniformes, continues et nulles à l'infini, sont les « valeurs propres » de l'équation (1). A une valeur propre W_n correspond donc au moins un ensemble [1] de fonctions propres $\Psi_{1,n}\Psi_{2,n}\Psi_{3,n}$

[1] L'ensemble des 4 fonctions propres $\Psi_{k,n}$ peut être nommée une « solution propre » de l'équation (1).

et $\Psi_{4,n}$. Nous donnons deux indices à chacun des Ψ_k, le premier étant l'indice introduit par la théorie de Dirac, le second celui qui caractérise la valeur propre correspondante.

Les quatre fonctions $\Psi_{k,n}$ correspondant à la valeur propre W_n satisfont donc par définition aux équations :

$$\left[\left(\frac{W_n}{c}+\frac{eV}{c}\right)+\sum_1^3{}_i \alpha_i P_i + \alpha_4 m_0 c\right] \Psi_{k,n} = 0 \ ; \ k = 1, 2, 3, 4. \quad (2)$$

Comme on peut multiplier les quatre fonctions $\Psi_{k,n}$ par une même constante arbitraire sans qu'elles cessent de satisfaire aux équations (2), on voit que chaque solution propre n'est définie qu'à une constante multiplicative complexe près. On détermine le module de cette constante à l'aide de la condition de normalisation qui, nous l'avons vu, doit s'écrire :

$$\int \sum_1^4{}_k \Psi^*_{k,n} \Psi_{k,n} \, d\tau = 1 \, . \quad (3)$$

Nous allons maintenant démontrer 1° que les valeurs propres W_n sont toutes réelles, 2° que si les fonctions propres $\Psi_{k,n}$ et $\Psi_{k,m}$ correspondent à des valeurs propres W_n et W_m distinctes, on a la condition d'orthogonalité :

$$\int \sum_1^4{}_k \Psi^*_{k,m} \Psi_{k,n} \, d\tau = 0 \, . \quad (4)$$

En effet, les fonctions $\Psi_{k,n}$ obéissant aux équations (2), les fonctions $\Psi^*_{k,m}$ obéissent aux équations :

$$\left[\left(\frac{W_m^*}{c}+\frac{eV}{c}\right)+\sum_1^3{}_i \alpha_i^* P_i^* + \alpha_4^* m_0 c\right] \Psi^*_{k,m} = 0 \ (k=1, 2, 3, 4). \quad (5)$$

Multiplions (2) par $\Psi^*_{k,m}$ et (5) par $\Psi_{k,n}$; sommons sur l'indice k et retranchons ; il vient :

$$\sum_{1}^{4}{}_k \left[\frac{W_n - W_m^*}{c} \Psi^*_{k,m} \Psi_{k,n} + \Psi^*_{k,m} \sum_{1}^{3}{}_i \alpha_i P_i \cdot \Psi_{k,n} - \right.$$

$$- \Psi_{k,n} \sum_{1}^{3}{}_i \alpha_i^* P_i^* \cdot \Psi^*_{k,m} + m_0 c (\Psi^*_{k,m} \alpha_4 \Psi_{k,n} +$$

$$\left. + \Psi_{k,n} \alpha^*_4 \Psi^*_{k,m}) \right] = 0 . \qquad (6)$$

Or nous allons montrer que, si F est un opérateur linéaire opérant sur les coordonnées (et non sur les indices k de Dirac), on a :

$$\sum_{1}^{4}{}_k \Psi^*_{k,m} \cdot F\alpha_i \Psi_{k,n} = \sum_{1}^{4}{}_k F\Psi_{k,n} \cdot \alpha_i^* \Psi^*_{k,m} (i = 1, 2, 3) . \quad (7)$$

En effet, en tenant compte de l'hermiticité des α_i, on trouve :

$$\sum_{1}^{4}{}_k \Psi^*_{k,m} F\alpha_i \Psi_{k,n} =$$

$$= \sum_{1}^{4}{}_k \Psi^*_{k,m} \cdot F \sum_{1}^{4}{}_j (\alpha_i)_{k,j} \Psi_{j,n} = \sum_{1}^{4}{}_{j,k} \Psi^*_{k,m} (\alpha_i)_{k,j} F\Psi_{j,n}$$

$$= \sum_{1}^{4}{}_{j,k} \Psi^*_{k,m} (\alpha_i^*)_{j,k} F\Psi_{j,n} = \sum_{1}^{4}{}_j F\Psi_{j,n} \cdot \sum_{1}^{4}{}_k (\alpha_i^*)_{j,k} \Psi^*_{k,m}$$

$$= \sum_{1}^{4}{}_j F \Psi_{j,n} \cdot \alpha_i^* \Psi^*_{j,m} \qquad \text{C. Q. F. D.} \qquad (8)$$

De (7) nous tirons d'abord comme cas particulier en faisant $F = 1$:

$$\sum_{1}^{4}{}_k \Psi^*_{k,m} \alpha_i \Psi_{k,n} = \sum_{1}^{4}{}_k \Psi_{k,n} \cdot \alpha_i^* \Psi^*_{k,m}. \qquad (9)$$

Il en résulte que dans (6) les termes en eA_x, eA_y, eA_z et en $m_0 c$ disparaissent. L'équation (6) se réduit donc à :

$$\sum_{1}^{4}{}_k \left[\frac{W_n - W_m^*}{c} \Psi^*_{k,m} \Psi_{k,n} - \right.$$

$$- \frac{h}{2\pi i} \Psi^*_{k,m} \left(\frac{\partial}{\partial x} \alpha_1 + \frac{\partial}{\partial y} \alpha_2 + \frac{\partial}{\partial z} \alpha_3\right) \Psi_{k,n}$$

$$-\frac{h}{2\pi i}\Psi_{k,n}\left(\frac{\partial}{\partial x}\alpha_1^* + \frac{\partial}{\partial y}\alpha_2^* + \frac{\partial}{\partial z}\alpha_3^*\right)\Psi^*_{k,m}\Bigg] = 0. \qquad (10)$$

Or d'après (7) on a :

$$\sum_1^4{}_k \Psi^*_{k,m}\frac{\partial}{\partial x}\alpha_1\Psi_{k,n} = \sum_1^4{}_k \frac{\partial\Psi_{k,n}}{\partial x}\cdot\alpha_1^*\Psi^*_{k,m} \text{ etc.} \qquad (11)$$

et la formule (10) devient :

$$\sum_1^4{}_k \left[\frac{W_n - W_m^*}{c}\Psi^*_{k,m}\cdot\Psi_{k,n} - \frac{h}{2\pi i}\left\{\frac{\partial}{\partial x}(\Psi_{k,n}\cdot\alpha_1^*\Psi^*_{k,m}) + \right.\right.$$
$$\left.\left. + \frac{\partial}{\partial y}(\Psi_{k,n}\cdot\alpha_2^*\Psi^*_{k,m}) + \frac{\partial}{\partial z}(\Psi_{k,n}\alpha_3^*\Psi^*_{k,m})\right\}\right] = 0. \qquad (12)$$

En intégrant dans tout l'espace, le terme entre accolades donne zéro parce que les Ψ_k sont nulles à l'infini et il reste

$$\frac{W_n - W_m^*}{c}\int\sum_1^4{}_k \Psi^*_{k,m}\Psi_{k,n}\,d\tau = 0. \qquad (13)$$

Si l'on fait d'abord $n = m$, on déduit de (13) que $W_n^* = W_n$ pour toute valeur de n et cela signifie que tous les W_n sont réels. Si ensuite on prend $n \neq m$ en se souvenant que par hypothèse $W_n \neq W_m$ on déduit de (13) la formule d'orthogonalité (4). Le théorème ainsi annoncé est donc démontré.

La relation d'orthogonalité est en général en défaut pour deux solutions propres correspondant à une même valeur propre (cas de la dégénérescence). Mais alors les solutions propres qui correspondent à une même valeur propre n'étant déterminées qu'à une transformation linéaire près, on peut toujours choisir ces solutions propres de façon qu'elles soient orthogonales.

Le cas des spectres continus de valeurs propres donnerait lieu aux mêmes remarques que dans la Mécanique ondulatoire à une fonction Ψ. D'une façon générale d'ailleurs, le parallélisme des deux théories est ici complet. Néanmoins il faut noter une différence : dans les formules où intervient une intégration dans l'espace telles que les formules de nor-

malisation et d'orthogonalité (3) et (4), il faut en théorie de Dirac faire une sommation sur l'indice k. La présence de cette sommation paraîtra toute naturelle quand nous développerons le point de vue synthétique dans lequel on considère l'indice k comme une sorte de variable supplémentaire discontinue (cf. chapitre XVI).

2. ***Matrices et intégrales premières dans la théorie de Dirac.*** — Dans la Mécanique ondulatoire à une fonction Ψ, on fait correspondre à tout opérateur linéaire et hermitique A la matrice dont les éléments sont définis par la formule :

$$A_{mn} = \int_D \Psi_m^* A (\Psi_n) \, d\tau \qquad (14)$$

$d\tau$ étant l'élément de volume du domaine d'existence D des fonctions Ψ_n. On dit alors que l'opérateur A est intégrale première pour le problème considéré si tous les A_{mn} sont indépendants du temps et nous avons vu qu'en posant

$$L = H - \frac{h}{2\pi i} \cdot \frac{\partial}{\partial t},$$

la condition nécessaire et suffisante pour que A soit intégrale première est :

$$LA - AL \equiv 0 \qquad (15)$$

Comment allons-nous transposer ces définitions dans la théorie de Dirac ?

Pour définir les éléments de matrice, nous tiendrons compte de la remarque faite à la fin du dernier paragraphe, c'est-à-dire que nous superposerons à l'intégration figurant dans l'ancienne définition (14) une sommation sur l'indice des fonctions Ψ_k de Dirac. Les éléments de la matrice correspondant à un opérateur [1] linéaire et hermitique A seront donc définis par la relation :

$$A_{mn} = \int_D \sum_{k=1}^{4} \Psi^*_{k,m} A (\Psi_{k,n}) \, d\tau. \qquad (16)$$

[1] Naturellement en théorie de Dirac, un opérateur peut opérer aussi bien sur l'indice k que sur les coordonnées.

Nous dirons toujours que A est intégrale première si tous les A_{mn} sont indépendants du temps. Pour trouver la condition qui exprime que A est intégrale première, nous écrirons l'équation symbolique de Dirac sous la forme condensée:

$$L(\Psi)=0 \tag{17}$$

avec :

$$L=\left(P_4+\sum_1^3{}_i\,\alpha_i\,P_i+\alpha_4\,m_0c\right) \tag{18}$$

et nous remarquerons qu'en posant :

$$H=-\left[e\,V+\sum_1^3{}_i\,c\alpha_i\,P_i+\alpha_4 m_0\,c^2\right] \tag{19}$$

on peut écrire :

$$L=\frac{1}{c}\left(-H+\frac{h}{2\pi i}\frac{\partial}{\partial t}\right);\qquad \frac{h}{2\pi i}\frac{\partial\Psi}{\partial t}=H(\Psi)\,. \tag{20}$$

L'opérateur H défini par (19) est l'opérateur Hamiltonien dans la théorie de Dirac. L'opérateur H est hermitique c'est-à-dire qu'il satisfait à la condition :

$$\int_D\sum_1^4{}_k\,\Psi^*_{km}\,H(\Psi_{k,n})d\tau=\int_D\sum_1^4{}_k\,\Psi_{kn}\,H^*\,(\Psi^*_{k,m})\,d\tau \tag{21}$$

pour toutes valeurs de m et de n. L'équation (21) est la généralisation évidente de la condition d'hermiticité de la Mécanique à un Ψ ; elle exprime d'ailleurs simplement que l'on a $H_{mn}=H^*_{nm}$ d'après la définition (16). La formule (21) se démontre aisément en tenant compte de l'hermiticité des α_i et de la propriété des $\Psi_{k,n}$ d'être nulles aux limites de D [1].

---

[1] Pour faire la démonstration, on remplace dans (21) l'opérateur H par son expression (19) et on vérifie que la relation (21) est valable pour chaque terme. Pour les termes contenant V, la vérification est immédiate. Pour les termes contenant A_x, A_y et A_z, on a par exemple :

$$\sum_1^4{}_k\,e\,A_x\sum_1^4{}_j(\alpha_1{}^*)_{kj}\,\Psi_{k,n}=\sum_1^4{}_j e\,A_x\,\Psi^*_{j,m}\sum_1^4{}_k\,(\alpha_1)_{j.k}\,\Psi_{k,n}=\sum_1^4{}_j\,\Psi^*_{j,m}eA_x\alpha_1\,\Psi_{j.m}$$

Pour exprimer que A est intégrale première, nous écrirons :

$$\frac{\partial A_{mn}}{\partial t} = \int_D \sum_{1}^{4}{}_k \left[\Psi^*_{k,m} A \left(\frac{\partial \Psi_{k,n}}{\partial t} \right) + \right.$$

$$\left. + \Psi^*_{k,m} \frac{\partial A}{\partial t} (\Psi_{k,m}) + \frac{\partial \Psi^*_{k,m}}{\partial t} A (\Psi_{k,n}) \right] d\tau = 0. \qquad (22)$$

Puis, comme l'on a $L(\Psi_{k,n}) = 0$ et $L^*(\Psi^*_{k,m}) = 0$, nous remplacerons les dérivées $\frac{\partial \Psi_{k,n}}{\partial t}$ et $\frac{\partial \Psi^*_{k,m}}{\partial t}$ par $\frac{2\pi i}{h} H(\Psi_{k,n})$ et $-\frac{2\pi i}{h} H^*(\Psi^*_{k,m})$ respectivement, ce qui nous donnera :

$$\frac{dA_{mn}}{dt} = \int_D \sum_{1}^{4}{}_k \left[\frac{2\pi i}{h} \Psi^*_{km} A H (\Psi_{k,n}) + \right.$$

$$\left. + \Psi^*_{km} \frac{\partial A}{\partial t} (\Psi_{k,n}) - \frac{2\pi i}{h} H^* (\Psi^*_{k,m}) A (\Psi_{k,n}) \right] d\tau. \qquad (23)$$

On peut transformer cette expression à l'aide de la formule:

$$\int_D \sum_{1}^{4}{}_k H^* (\Psi^*_{k,m}) A (\Psi_{k,n}) d\tau = \int_D \sum_{1}^{4}{}_k \Psi^*_{k,m} HA (\Psi_{k,n}) d\tau \quad (24)$$

à cause de l'hermiticité des α_i, et la relation (21) est ainsi démontrée pour ces termes. Restent les termes tels que

$$\sum_{1}^{4}{}_k \frac{h}{2\pi i} \frac{\partial}{\partial x} \alpha_1^* \Psi^*_{k,m} \cdot \Psi_{kn} .$$

On a d'abord, toujours à cause de l'hermiticité des α_i :

$$\frac{h}{2\pi i} \sum_{1}{}_{k,j} \frac{\partial}{\partial x} \left[(\alpha_1^*)_{k,j} \Psi^*_{j,m} \right] \Psi_{k,n} = \frac{h}{2\pi i} \sum_{1}^{4}{}_j \frac{\partial \Psi^*_{j,m}}{\partial x} \sum_{1}^{4}{}_k (\alpha_1)_{j,k} \Psi_{k,n}$$

puis par intégration par parties parce que les $\Psi_{k,n}$ sont nuls aux limites de D :

$$\frac{h}{2\pi i} \int_D \sum_{1}^{4}{}_j \frac{\partial \Psi^*_{j,m}}{\partial x} \sum_{1}^{4}{}_k (\alpha_1)_{j,k} \Psi_{k,n} d\tau = -$$

$$-\frac{h}{2\pi i} \int_D \sum_{1}^{4}{}_j \Psi^*_{j,m} \sum_{1}^{4}{}_k (\alpha_1)_{j,k} \frac{\partial \Psi_{k,n}}{\partial x} = \int_D \sum_{1}^{4}{}_k \Psi^*_{k,m} \left(-\frac{h}{2\pi i} \frac{\partial}{\partial x} \right) \alpha_1 \Psi_{k,n}$$

C. Q. F. D.

qui se démontre comme (21). En portant (24) dans (23), on trouve :

$$\frac{dA_{m,n}}{dt} = \int_{D} \sum_{1}^{4} {}_{k} \Psi^{*}_{k,m} \left[\frac{\partial A}{\partial t} + \frac{2\pi i}{h} (AH - HA) \right] \Psi_{k,n} \, d\tau. \quad (25)$$

On en conclut d'après (22) que la condition nécessaire et suffisante pour que A soit intégrale première est :

$$\frac{\partial A}{\partial t} + \frac{2\pi i}{h} (AH - HA) = 0 \quad (26)$$

car les $\Psi_{k,n}$ forment un système complet.

On peut encore écrire la condition (26) autrement en remarquant que l'on a pour une fonction f quelconque :

$$\frac{\partial}{\partial t} A(f) = \frac{\partial A}{\partial t}(f) + A\left(\frac{\partial f}{\partial t}\right) \quad (27)$$

d'où symboliquement :

$$\frac{\partial A}{\partial t} = \frac{\partial}{\partial t} \cdot A - A \cdot \frac{\partial}{\partial t}. \quad (28)$$

Donc en remplaçant l'opérateur $\frac{\partial A}{\partial t}$ dans (26) par sa valeur (28), on a :

$$LA - AL \equiv 0 \quad (29)$$

en vertu de (20).

Les conditions (26) et (29) ont la même forme que les conditions analogues de la Mécanique à un Ψ, mais avec une définition différente des opérateurs L et H.

Nous remarquerons que l'équation (1) peut s'écrire en introduisant la définition (19) de H sous la forme :

$$H(\Psi_k) = W\Psi_k. \quad (30)$$

Les valeurs propres W_k définies au paragraphe 1 sont donc bien les valeurs propres de l'opérateur Hamiltonien, c'est-à-dire les valeurs propres de l'énergie.

3. ***Exemples d'intégrales premières. Moment de rotation propre de l'électron.*** — Nous allons d'abord examiner dans quel cas l'opérateur Hamiltonien correspondant à l'énergie est une intégrale première. Pour cela il faut que l'on ait :

$$\mathrm{LH} - \mathrm{HL} = \frac{1}{c}\left(-\mathrm{H} + \frac{h}{2\pi i}\frac{\partial}{\partial t}\right)\mathrm{H} - \mathrm{H}\frac{1}{c}\left(-\mathrm{H} + \frac{h}{2\pi i}\frac{\partial}{\partial t}\right) \equiv 0 \tag{31}$$

ou plus simplement

$$\frac{\partial}{\partial t}\ \mathrm{H} - \mathrm{H}\,.\frac{\partial}{\partial t} = \frac{\partial \mathrm{H}}{\partial t} \equiv 0. \tag{32}$$

La condition nécessaire et suffisante pour que l'énergie soit intégrale première est donc toujours que le champ extérieur (défini ici par les 4 fonctions V, A_x, A_y, A_z) soit indépendant du temps. C'est le théorème de la conservation de l'énergie en Mécanique de Dirac.

On voit aussi aisément que si le potentiel vecteur est nul et si le potentiel scalaire ne dépend pas d'une des coordonnées, mettons x, la composante correspondante de la quantité de mouvement $\left(p_x = -\frac{h}{2\pi i}\frac{\partial}{\partial x}\right)$ est intégrale première. C'est le théorème de la conservation de la quantité de mouvement.

Beaucoup plus intéressante est l'étude du théorème de la conservation du moment de la quantité de mouvement dans la théorie de Dirac. Nous avons vu en effet que dans la Mécanique ondulatoire à une fonction Ψ, le moment de rotation autour de l'axe des z correspondant à l'opérateur :

$$\mathrm{M}_z = x\,p_y - y\,p_x = \frac{h}{2\pi i}\left(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\right) \tag{33}$$

est une intégrale première si le champ de force possède la symétrie cylindrique autour de l'axe des z. Or nous allons voir que ce résultat ne subsiste pas dans la Mécanique à 4 fonctions Ψ et nous chercherons par quoi il doit être remplacé.

Vérifions d'abord que dans un champ dérivant d'un potentiel scalaire à symétrie cylindrique $V(\rho, z)$ avec

$$\rho = \sqrt{x^2 + y^2},$$

M_z n'est pas une intégrale première en Mécanique de Dirac. Pour cela nous devons montrer que M_z ne permute pas avec L. Il est d'abord évident que M_z permute avec $\frac{h}{2\pi i}\frac{\partial}{\partial t}$, $\alpha_3 \frac{h}{2\pi i}\frac{\partial}{\partial z}$ et $\alpha_4 m_0 c$; il permute aussi avec le terme $\frac{e}{c}V$ car l'on a (en opérateurs) :

$$V\left(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\right) - \left(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\right)V \equiv x\frac{\partial V}{\partial y} - y\frac{\partial V}{\partial x} \tag{34}$$

ce qui est nul parce que V ne dépend par hypothèse que de

$$\rho = \sqrt{x^2 + y^2}.$$

Mais par contre M_z ne permute pas avec les termes en

$-\alpha_1 \frac{h}{2\pi i}\frac{\partial}{\partial x}$ et en $-\alpha_2 \frac{h}{2\pi i}\frac{\partial}{\partial y}$, car on a :

$$\left(-\alpha_1 \frac{h}{2\pi i}\frac{\partial}{\partial x} - \alpha_2 \frac{h}{2\pi i}\frac{\partial}{\partial y}\right)\frac{h}{2\pi i}\left(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\right) -$$

$$- \frac{h}{2\pi i}\left(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\right)\left(-\alpha_1 \frac{h}{2\pi i}\frac{\partial}{\partial x} - \alpha_2 \frac{h}{2\pi i}\frac{\partial}{\partial y}\right)$$

$$= \frac{h^2}{4\pi^2}\left(\alpha_2\frac{\partial}{\partial x} - \alpha_1\frac{\partial}{\partial y}\right) \tag{35}$$

On a donc au total :

$$L M_z - M_z L = \frac{h^2}{4\pi^2}\left(\alpha_2\frac{\partial}{\partial x} - \alpha_1\frac{\partial}{\partial y}\right) \tag{36}$$

et M_z n'est pas intégrale première.

Considérons alors l'opérateur :

$$N_z = M_z - \alpha_1 \alpha_2 \frac{h}{4\pi i}. \tag{37}$$

Cet opérateur est hermitique car le produit des deux matrices α_1 et α_2 qui sont hermitiques et anticommutent, est antihermitique et leur quotient par i est hermitique. Nous allons montrer que dans le cas considéré, N_z est intégrale première. Pour cela, formons la différence

$$L\,\alpha_1\,\alpha_2\frac{h}{4\pi i}-\alpha_1\,\alpha_2\frac{h}{4\pi i}L\,.$$

Les termes en $\dfrac{\partial}{\partial t}$, V et $\alpha_4 m_0 c$ de L commutent évidemment avec $\alpha_1\alpha_2$; il en est de même du terme $\alpha_3\dfrac{h}{2\pi i}\dfrac{\partial}{\partial z}$ en vertu des propriétés des α_i. Il reste donc :

$$L\,\alpha_1\,\alpha_2\frac{h}{4\pi i}-\alpha_1\,\alpha_2\frac{h}{4\pi i}L=$$
$$\frac{h^2}{8\pi^2}\left[\alpha_1\,\alpha_1\,\alpha_2\frac{\partial}{\partial x}+\alpha_2\,\alpha_1\,\alpha_2\frac{\partial}{\partial y}-\alpha_1\,\alpha_2\,\alpha_1\frac{\partial}{\partial x}-\alpha_1\,\alpha_2\,\alpha_2\frac{\partial}{\partial y}\right].\qquad(38)$$

Comme on a $\alpha_1\,\alpha_1=\alpha_2\,\alpha_2=1$, $\alpha_2\,\alpha_1\,\alpha_2=-\alpha_1$ et $\alpha_1\,\alpha_2\,\alpha_1=-\alpha_2$, on trouve :

$$L\,\alpha_1\,\alpha_2\frac{h}{4\pi i}-\alpha_1\,\alpha_2\frac{h}{4\pi i}L=\frac{h^2}{4\pi^2}\left(\alpha_2\frac{\partial}{\partial x}-\alpha_1\frac{\partial}{\partial y}\right)=LM_z-M_zL.\qquad(39)$$

On en conclut la relation

$$L\,N_z-N_z\,L\equiv 0\qquad(40)$$

et l'on voit que N_z est intégrale première.

On peut faire un raisonnement analogue au précédent pour les composantes d'indices x et y en permutant le rôle des axes. On voit alors qu'en théorie de Dirac, le moment de rotation *total* de l'électron est un vecteur de composantes :

$$N_x=M_x+S_x\quad N_y=M_y+S_y\quad N_z=M_z+S_z\qquad(41)$$

les opérateurs correspondant M_x etc. étant :

$$M_x=\frac{h}{2\pi i}\left(z\frac{\partial}{\partial y}-y\frac{\partial}{\partial z}\right);\ M_y=\frac{h}{2\pi i}\left(x\frac{\partial}{\partial z}-z\frac{\partial}{\partial x}\right);$$
$$M_z=\frac{h}{2\pi i}\left(y\frac{\partial}{\partial x}-x\frac{\partial}{\partial y}\right).\qquad(42)$$

$$S_x = -\alpha_2\,\alpha_3 \frac{h}{4\pi i}\,;\; S_y = -\alpha_3\,\alpha_1 \frac{h}{4\pi i}\,;\; S_z = -\alpha_1\,\alpha_2 \frac{h}{4\pi i}\,.$$

$M_x\,M_y\,M_z$ sont les composantes du moment de rotation « orbital » de l'électron et l'on est amené tout naturellement à considérer S_x, S_y, S_z comme les composantes du moment de rotation propre de l'électron, c'est-à-dire du « spin ».

4. *Calcul explicite de N_z. Signe de la masse propre pour les ondes Ψ_k.* — Il est très intéressant de calculer explicitement l'opérateur N_z correspondant à l'axe privilégié z et de voir quels résultats on obtient en l'appliquant successivement aux quatre Ψ_k.

En partant des valeurs que nous connaissons de α_1 et α_2, on calcule aisément à l'aide de la règle de multiplication des matrices la matrice produit $\alpha_1\,\alpha_2$ et l'on trouve :

$$\alpha_1\,\alpha_2 = \begin{vmatrix} -i & 0 & 0 & 0 \\ 0 & +i & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & 0 & 0 & +i \end{vmatrix} \tag{43}$$

De là, on tire facilement :

$$\begin{aligned} N_z(\Psi_1) &= M_z(\Psi_1) + \frac{h}{4\pi}\Psi_1 \\ N_z(\Psi_2) &= M_z(\Psi_2) - \frac{h}{4\pi}\Psi_2 \\ N_z(\Psi_3) &= M_z(\Psi_3) + \frac{h}{4\pi}\Psi_3 \\ N_z(\Psi_4) &= M_z(\Psi_4) - \frac{h}{4\pi}\Psi_4\,. \end{aligned} \tag{44}$$

Nous voyons donc sur les formules (44) que pour les ondes d'indice impair le moment de rotation propre est $+\frac{h}{4\pi}$ et pour les ondes d'indice pair $-\frac{h}{4\pi}$.

Ecrivons la valeur moyenne de N_z. Nous avons :

$$\overline{N_z} = \iiint \sum_{1}^{4}{}_k \Psi_k^* \, N_z(\Psi_k) \, d\tau$$
$$= \overline{M}_z + \frac{h}{4\pi} \iiint [\Psi_1^* \Psi_1 - \Psi_2^* \Psi_2 + \Psi_3^* \Psi_3 - \Psi_4^* \Psi_4] \, d\tau. \tag{45}$$

On doit interpréter cette formule en disant : le moment de rotation propre de l'électron le long de l'axe des z ne peut prendre que les deux valeurs $\frac{h}{4\pi}$ et $-\frac{h}{4\pi}$, la probabilité de la première hypothèse étant

$$\iiint (\Psi_1^* \Psi_1 + \Psi_3^* \Psi_3) \, d\tau$$

et celle de la seconde étant

$$\iiint [\Psi_2^* \Psi_2 + \Psi_4^* \Psi_4] \, d\tau.$$

Comparons alors la formule (45) avec la formule (16) du chapitre XIV qui donne le moment magnétique moyen de l'électron le long de oz :

$$\overline{M}_z = \frac{e h}{4\pi m_0 c} \iiint (\Psi_1^* \Psi_1 - \Psi_2^* \Psi_2 - \Psi_3^* \Psi_3 + \Psi_4^* \Psi_4) \, d\tau. \tag{16}$$

Nous voyons qu'on peut faire correspondre :

à l'onde Ψ_3 un moment magnétique $-\frac{e h}{4\pi m_0 c}$

un moment de rotation $+\frac{h}{4\pi}$

à l'onde Ψ_4 un moment magnétique $+\frac{e h}{4\pi m_0 c}$

un moment de rotation $-\frac{h}{4\pi}$.

Pour ces deux ondes qui sont prépondérantes à l'approximation Newtonienne, nous l'avons vu, le rapport des deux

moments est égal à $-\frac{e}{m_0 c}$ comme on devait s'y attendre à cause du double magnétisme de l'électron tournant (1).

Mais pour les ondes Ψ_1 et Ψ_2, nous obtenons un résultat surprenant. En effet d'après (45) et (46), il faut faire correspondre :

à l'onde Ψ_1 un moment magnétique $+\frac{eh}{4\pi m_0 c}$

un moment de rotation $+\frac{h}{4\pi}$

à l'onde Ψ_2 un moment magnétique $-\frac{eh}{4\pi m_0 c}$

un moment de rotation $-\frac{h}{4\pi}$

d'où pour le rapport des deux moments $+\frac{e}{m_0 c}$, valeur qui diffère par son signe de la valeur attendue.

D'où provient cette anomalie? Pour le comprendre, remarquons d'abord que tout se passe comme si, en ce qui concerne les ondes Ψ_1 et Ψ_2, la masse propre de l'électron était $-m_0$ au lieu de m_0. Ceci se voit bien sur les équations mêmes de Dirac :

$$\left(P_4 + \sum_1^4{}_i \alpha_i P_i + \alpha_4 m_0 c\right)\Psi_k = 0 \qquad (k = 1, 2, 3, 4) \quad (47)$$

si l'on retient la forme explicite de la matrice α_4 :

$$\alpha_4 = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{vmatrix} \quad (48)$$

Le terme en m_0 a un signe différent dans les deux premières

(1) Le rapport $-\frac{e}{m_0 c}$ est en effet le double du rapport normal donné par la formule (7) du Chapitre IV.

équations (47) d'une part, dans les deux dernières d'autre part.

On peut exprimer ce fait en disant qu'en théorie de Dirac, la masse propre m_0 de l'ancienne Mécanique est remplacée par l'opérateur $-\alpha_4 m_0$. Comme l'on a :

$$-\alpha_4 m_0 \Psi_1 = -m_0 \Psi_1 \,; \quad -\alpha_4 m_0 c \Psi_2 = -m_0 \Psi_2 \,;$$
$$-\alpha_4 m_0 c \Psi_3 = m_0 \Psi_3 \,; \quad -\alpha_4 m_0 c \Psi_4 = m_0 \Psi_4 \qquad (49)$$

il est visible qu'aux ondes Ψ_3 et Ψ_4 correspond la masse m_0 tandis qu'aux ondes Ψ_1 et Ψ_2 correspond la masse $-m_0$. Si le lecteur veut bien maintenant revoir le raisonnement qui précède la formule (32) du chapitre X, il en comprendra mieux la portée à la lumière de ce qui vient d'être dit.

CHAPITRE XVI

Résumé systématique des résultats obtenus

1. *L'indice des fonctions Ψ_k considéré comme une variable.*

JUSQU'ICI nous avons exposé la théorie de Dirac en disant qu'elle comporte l'existence de 4 fonctions d'ondes Ψ_1, Ψ_2, Ψ_3, Ψ_4 des 4 variables $xyzt$. On peut se placer à un autre point de vue et dire qu'il y a une *seule* fonction Ψ dépendant des 4 variables continues $xyzt$ qui sont susceptibles de prendre toutes les valeurs réelles de $-\infty$ à $+\infty$ et d'une cinquième variable ζ, la variable de « spin », ne pouvant prendre que les 4 valeurs : 1, 2, 3, 4. Cela revient à considérer l'indice k des fonctions Ψ_k comme une variable discontinue à 4 valeurs possibles.

Les opérateurs α_i opèrent sur la variable discontinue ζ tandis que les opérateurs comme p_x par exemple opèrent sur les variables continues. Il est loisible de considérer des opérateurs qui opèrent à la fois sur les 4 variables (1) $xzy\zeta$: tel est l'Hamiltonien H de Dirac défini par la formule (19) du dernier chapitre.

(1) Bien que la théorie de Dirac soit en un sens relativiste, le temps y joue un rôle différent des quatre autres variantes $x\,y\,z\,\zeta$; nous reviendrons plus tard sur ce point.

Écrivons maintenant l'équation de Dirac sous la forme :

$$\left(P_4 + \sum_1^3 {}_i \alpha_i P_i + \alpha_4 m_0 c\right) \Psi(x, y, z, t, \zeta) = 0 \qquad (1)$$

Jusqu'à présent, nous avons regardé cette équation comme une équation symbolique résumant 4 équations distinctes. Nous pouvons maintenant la considérer comme une équation unique de propagation au cours du temps dans l'espace des 4 variables $xyz\zeta$.

Les intégrations que l'on effectuait dans la Mécanique ondulatoire « sans spin » dans l'espace à trois dimensions xyz, on doit ici les effectuer dans l'espace $xyz\zeta$. Ceci nous explique la raison profonde du fait suivant déjà signalé précédemment : toutes les formules de la Mécanique ondulatoire où figure une intégration dans l'espace doivent être complétées en théorie de Dirac par une sommation de 1 à 4 sur l'indice k

Cette sommation $\sum_1^4 {}_k$ correspond en effet à une sorte d'intégration sur la variable discontinue ζ et il nous sera souvent commode de la représenter symboliquement par une intégrale $\int \cdots d\zeta$. Ainsi par exemple, la condition de normalisation de la Mécanique ondulatoire sans spin :

$$\int\int\int_{-\infty}^{+\infty} \Psi^*(x, y, z, t) \cdot \Psi(x, y, z, t)\, dx\, dy\, dz = 1 \qquad (2)$$

devient en théorie de Dirac :

$$\int\int\int_{-\infty}^{+\infty} \sum_1^4 {}_k \Psi_k^* \cdot \Psi_k\, dx\, dy\, dz =$$

$$\int\int\int\int \Psi^*(x, y, z, t, \zeta) \cdot \Psi(x, y, z, t, \zeta)\, dx\, dy\, dz\, d\zeta = 1. \quad (3)$$

De même, le nouveau point de vue permet de voir aisément comment on doit écrire en théorie de Dirac les dévelop-

pements en série de fonctions propres. Soit en effet une fonction $f(x, y, z, t, \zeta)$ des cinq variables x, y, z, t, ζ. Elle est équivalente à quatre fonctions f_1, f_2, f_3, f_4 des quatre variables continues x, y, z, t. Supposons que nous connaissions le système de fonctions propres de l'opérateur Hamiltonien ; nous avons déjà convenu de désigner une fonction propre de ce système par l'ensemble des 4 fonctions $\Psi_{1,m}$ $\Psi_{2,m}$ $\Psi_{3,m}$ $\Psi_{4,m}$ où le second indice m caractérise la valeur propre correspondante. A notre nouveau point de vue, nous devons représenter l'ensemble des 4 fonctions $\Psi_{k,m}$ par la fonction unique : $\Psi_m(x, y, z, t, \zeta)$.

Le développement de la fonction $f(x, y, z, t, \zeta)$ suivant le système complet des $\Psi_m(x, y, z, t, \zeta)$ s'écrira alors tout naturellement :

$$f(x, y, z, t, \zeta) = \sum_m c_m \Psi_m (x, y, z, t, \zeta) \tag{4}$$

ce qui équivaut aux quatre relations :

$$f_k (x, y, z, t) = \sum_m c_m \Psi_{k,m} (x, y, z, t) . \quad (k = 1, 2, 3, 4) \tag{5}$$

On voit donc que chacune des 4 composantes f_k se développe suivant les fonctions propres $\Psi_{k,m}$ *de même indice* k, les coefficients du développement étant en outre les mêmes pour les 4 composantes. Cette proposition qui peut ne pas paraître évidente quand on considère k comme un indice, semble au contraire toute naturelle quand on raisonne comme nous venons de le faire en remplaçant l'indice k par la variable discontinue ζ.

2. ***Enoncé des principes généraux dans la théorie de Dirac.*** — L'introduction de la variable discontinue ζ permet de trouver immédiatement comment les principes généraux de la Mécanique ondulatoire se transposent dans la théorie de Dirac.

Nous admettrons d'abord qu'à toute grandeur physique observable attaché à l'électron correspond un opérateur A

qui en général peut opérer sur les quatre variables x, y, z, ζ. Cet opérateur doit toujours être hermitique dans l'espace x, y, z, ζ, c'est-à-dire que l'on doit avoir :

$$\int\int\int\int f^* (x, y, z, t, \zeta) \, A \, g \, (x, y, z, t, \zeta) \, dx \, dy \, dz \, d\zeta =$$

$$= \int\int\int\int g \, (x, y, z, t, \zeta) \, A^* \, g^* \, (x, y, z, t, \zeta) \, dx \, dy \, dz \, d\zeta \quad (6)$$

La valeurs propres de l'opérateur A sont définies comme étant les valeurs de la constante a pour lesquelles l'équation :

$$A \, \varphi \, (x, y, z, t, \zeta) \equiv a \, \varphi \, (x, y, z, t, \zeta) \qquad (7)$$

admet au moins une solution partout finie, continue et uniforme en xyz pour chacune des 4 valeurs de ζ. Ces valeurs propres, généralement fonction du temps, sont réelles et deux fonctions propres φ_m et φ_n correspondant à deux valeurs propres distinctes a_m et a_n sont orthogonales dans l'espace x, y, z, ζ c'est-à-dire que l'on a :

$$\int\int\int\int \varphi_m^* \, (x, y, z, t, \zeta) \, \varphi_n \, (x, y, z, t, \zeta) \, dx \, dy \, dz \, d\zeta = 0 \quad (8)$$

On démontre ces résultats par la même méthode que nous avons employée au chapitre V, paragraphe 4, en ajoutant toujours aux intégrations en x, y, z une sommation en ζ. Nous supposerons de plus toujours que les fonctions propres φ_n ont été normées par la condition :

$$\int\int\int\int \varphi_n^* \, (x, y, z, t, \zeta) \, \varphi_n \, (x, y, z, t, \zeta) \, dx \, dy \, dz \, d\zeta = 1 \quad (9)$$

Ceci posé, il est facile d'énoncer en théorie de Dirac les principes fondamentaux de la nouvelle Mécanique.

Le premier de ces principes s'énoncera sans changement aucun en disant qu'une mesure de la grandeur A effectuée à l'instant t fournit nécessairement comme résultat une des valeurs propres de l'opérateur A à cet instant t.

Pour énoncer le second principe, nous supposerons d'abord que l'opérateur A est un opérateur « complet », c'est-à-dire qu'il intéresse toutes les 4 variables x, y, z, ζ ; de plus nous

le supposerons aussi non dégénéré c'est-à-dire n'ayant pas de valeurs propres multiples. Soit alors $\Psi(x, y, z, t, \zeta)$ la fonction d'onde d'un électron. Cette fonction d'onde peut se développer suivant les fonctions propres de l'opérateur A sous la forme

$$\Psi(x, y, z, \zeta, t) = \sum_m c_m \varphi_m(x, y, z, \zeta) \tag{10}$$

les c_m étant des constantes complexes, en général fonctions du temps. Le second principe affirme alors que $|c_m(t)|^2$ est la probabilité pour qu'une mesure de la grandeur A fournisse la valeur propre a_m correspondant à φ_m.

Si A possède des valeurs propres multiples, à un même a_m correspondent plusieurs φ_m. La probabilité de la valeur a_m pour la grandeur A est alors égale à la somme des carrés des modules des coefficients relatifs à ces φ_m dans le développement de Ψ.

Si l'opérateur A est incomplet c'est-à-dire n'intéresse que certaines des variables x, y, z, ζ, les φ_m correspondants ne dépendent que de ces variables et les coefficients c_m dans le développement de Ψ suivant les φ_m dépendent des variables qui ne figurent pas dans A. Il faut en ce cas, pour obtenir la probabilité de la valeur propre a_m, intégrer la quantité $|c_m|^2$ dans tout le domaine de celles des variables x, y, z, ζ dont c_m dépend. Par exemple, si A est un opérateur, comme l'un des α_i, agissant seulement sur la variable ζ, ces fonctions propres sont de la forme $\varphi_m(\zeta)$ et l'on a :

$$\Psi(x, y, z, t, \zeta) = \sum_m c_m(x, y, z, t)\, \varphi_m(\zeta) \tag{11}$$

La probabilité pour que la grandeur A ait la valeur a_m est alors donnée par

$$\iiint_{-\infty}^{+\infty} |c_m(x, y, z, t)|^2 \, dx\, dy\, dz\ .$$

Nous verrons une application de ce résultat au paragraphe suivant.

Ces principes admis, on voit que la valeur moyenne de la grandeur A est bien égale comme nous l'avions admis à :

$$\bar{A} = \int\int\int\int \Psi^* A(\Psi)\, dx\, dy\, dz\, d\zeta =$$

$$\int\int_{-\infty}^{+\infty}\int \sum_{1}^{4}{}_k \Psi_k^* A(\Psi_k)\, d\tau. \quad (12)$$

car si dans cette expression l'on remplace Ψ et Ψ^* par leurs développements suivant les fonctions propres de A, on voit aisément qu'elle est bien égale à la somme des produits de chaque valeur propre par sa probabilité. La quantité $\sum_{1}^{4}{}_k \Psi_k^* A(\Psi_k)$ peut donc être nommée « la densité de valeur moyenne » pour la grandeur A, mais comme nous l'avons signalé au chapitre VI, paragraphe 3, une telle densité ne peut pas être considérée comme ayant le même sens physique que les densités des théories classiques. Néanmoins, nous allons voir bientôt que les densités de ce genre correspondant aux opérateurs α_i et aux opérateurs hermitiques formés de produits des α_i sont réelles et possèdent un caractère tensoriel ce qui permet de les rapprocher de certaines grandeurs de la Physique classique.

Rappelons enfin que l'élément ij de la matrice correspondant à un opérateur A doit être défini en Mécanique de Dirac par la formule :

$$A_{ij} = \int\int\int\int \Psi_i^*(x, y, z, t, \zeta)\, A\Big(\Psi_j(x, y, z, t, \zeta)\Big)\, dx\, dy\, dz\, d\zeta$$

$$= \int\int_{-\infty}^{+\infty}\int \sum_{1}^{4}{}_k \Psi_{k,i}^* A(\Psi_{k,j})\, d\tau \quad (13)$$

et est égal au coefficient de $\Psi_i(x, y, z, t, \zeta)$ dans le développement de la fonction $A(\Psi_j)$ suivant les fonctions propres de l'opérateur Hamiltonien.

Nous avons implicitement supposé dans ce qui précède

que le spectre de l'opérateur A est purement discontinu. Ce que nous avons dit aux chapitres V et VI sur les spectres continus permettra au lecteur de voir sans peine comment les formules écrites ci-dessus sont à modifier s'il existe des valeurs propres formant une suite continue.

3. *Exemple: valeurs et fonctions propres de l'opérateur* $\mathcal{M}_z = \frac{eh}{4\pi m_0 c} i\alpha_1\alpha_2\alpha_4$. — Comme exemple d'un opérateur A agissant seulement sur la variable de spin, nous prendrons l'opérateur $\mathcal{M}_z = \frac{eh}{4\pi m_0 c} i\alpha_1\alpha_2\alpha_4$ que nous avons fait correspondre à la composante z du moment magnétique propre de l'électron (cf. chapitre X à la fin).

Pour abréger l'écriture nous poserons :

$$B = \frac{eh}{4\pi m_0 c} = \text{magnéton de Bohr.} \tag{14}$$

L'équation qui définit les valeurs et fonctions propres de l'opérateur $\mathcal{M}_z$ est, d'après (7) :

$$B\sum_1^4{}_k\, i(\alpha_1\alpha_2\alpha_4)_{jk}\,\varphi_j = a\,\varphi_j \quad (j = 1, 2, 3, 4). \tag{15}$$

En tenant compte de la valeur des éléments de la matrice $\alpha_1\alpha_2\alpha_4$, on trouve pour les quatre valeurs de la variable ζ:

$$B\varphi(1) = +a\varphi(1)\;;\; B\varphi(2) = -a\varphi(2)\;;$$
$$B\varphi(3) = -a\varphi(3)\;;\; B\varphi(4) = +a\varphi(4). \tag{16}$$

Ces équations n'admettent une solution non identiquement nulle $\varphi(\zeta)$ que si l'on a : $a = \pm$ B. L'opérateur $\mathcal{M}_z$ n'a donc pour valeurs propres que les valeurs $\pm$ B comme nous devions nous y attendre.

Pour la valeur propre $a = +$B, il existe deux fonctions propres *indépendantes* que nous nommerons

$$\varphi_+^{(1)}(\zeta) \text{ et } \varphi_+^{(2)}(\zeta).$$

Ces fonctions propres sont définies par leurs valeurs pour les 4 valeurs possibles de la variable ζ, savoir :

$$\begin{array}{llll} \varphi_+^{(1)}(1) = 1 & \varphi_+^{(1)}(2) = 0 & \varphi_+^{(1)}(3) = 0 & \varphi_+^{(1)}(4) = 0 \\ \varphi_+^{(2)}(1) = 0 & \varphi_+^{(2)}(2) = 0 & \varphi_+^{(2)}(3) = 0 & \varphi_+^{(2)}(4) = 1 \end{array} \qquad (17)$$

Ces fonctions sont normées et orthogonales car on a :

$$\int |\Psi_+^{(1)}|^2 d\zeta = 1 \quad \int |\Psi_+^{(2)}|^2 d\zeta = 1 \quad \int \Psi_+^{(1)*} \Psi_+^{(2)} d\zeta = 0 . \qquad (18)$$

La valeur propre $+B$ est donc double.

De même la valeur propre $a = -B$ est double car il lui correspond deux fonctions propres indépendantes, normées et orthogonales, $\varphi_-^{(1)}(\zeta)$ et $\varphi_-^{(2)}(\zeta)$, définies par :

$$\begin{array}{llll} \varphi_-^{(1)}(1) = 0 & \varphi_-^{(1)}(2) = 1 & \varphi_-^{(1)}(3) = 0 & \varphi_-^{(1)}(4) = 0 \\ \varphi_-^{(2)}(1) = 0 & \varphi_-^{(2)}(2) = 0 & \varphi_-^{(2)}(3) = 1 & \varphi_-^{(2)}(4) = 0. \end{array} \qquad (19)$$

On vérifie que les deux fonctions φ_- sont orthogonales aux deux fonctions φ_+.

Soit alors $\Psi(x, y, z, t, \zeta)$ la fonction d'onde d'un électron. En vertu du premier principe général, les valeurs possibles de la composante z de son moment magnétique propre sont les valeurs propres de l'opérateur $\mathcal{M}_z$, savoir $\pm$ B, c'est-à-dire ± 1 magnéton de Bohr. Pour trouver les probabilités respectives de ces deux hypothèses, nous devons écrire le développement :

$$\Psi(x,y,z,t,\zeta) = c_+^{(1)}(x,y,z,t)\,\varphi_+^{(1)}(\zeta) + c_+^{(2)}(x,y,z,t)\,\varphi_+^{(2)}(\zeta) + \\ + c_-^{(1)}(x,y,z,t)\,\varphi_-^{(1)}(\zeta) + c_-^{(2)}(x,y,z,t)\,\varphi_-^{(2)}(\zeta). \qquad (20)$$

D'après le second principe général, la probabilité de la valeur $+B$ est

$$\iiint_{-\infty}^{+\infty} \left[|c_+^{(1)}|^2 + |c_+^{(2)}|^2\right] d\tau$$

et celle de la valeur $-B$ est de même

$$\iiint_{-\infty}^{+\infty} \left[|c_-^{(1)}|^2 + |c_-^{(2)}|^2\right] d\tau$$

Or, en substituant successivement dans la formule (20) les valeurs $=1, 2, 3, 4$, on trouve aisément :

$$c_+^{(1)} = \Psi_1(x, y, z, t)\ ; \quad c_+^{(2)} = \Psi_4(x, y, z, t)\ ;$$
$$c_-^{(1)} = \Psi_2(x, y, z, t)\ ; \quad c_-^{(2)} = \Psi_3(x, y, z, t). \qquad (21)$$

La probabilité de la valeur $+$ B est donc

$$\int\!\!\int\!\!\int_{-\infty}^{+\infty} \left[|\Psi_1|^2 + |\Psi_4|^2\right] d\tau$$

et celle de la valeur de $-$ B est

$$\int\!\!\int\!\!\int_{-\infty}^{+\infty} \left[|\Psi_2|^2 + |\Psi_3|^2\right] d\tau.$$

C'est bien la conclusion à laquelle nous étions parvenus antérieurement.

L'exemple simple que nous venons de traiter montre comment on peut obtenir les valeurs et fonctions propres des opérateurs qui agissent seulement sur ζ. Ces considérations sont en particulier applicables aux opérateurs dont nous allons parler au début du paragraphe suivant, opérateurs qui ont tous pour valeurs propres doubles les valeurs ± 1.

4. *Les 16 opérateurs fondamentaux de la théorie de Dirac. Grandeurs et densités correspondantes.* — A l'aide des opérateurs α_i et de la matrice 1 (à 4 lignes et 4 colonnes) on peut former le tableau suivant contenant seize opérateurs hermitiques agissant seulement sur la variable ζ:

$$\begin{array}{c}
\alpha_4 \\
\alpha_1 \quad \alpha_2 \quad \alpha_3 \quad 1 \\
i\,\alpha_2\,\alpha_3\,\alpha_4 \quad i\,\alpha_3\,\alpha_1\,\alpha_4 \quad i\,\alpha_1\,\alpha_2\,\alpha_4 \quad i\,\alpha_1\,\alpha_4 \quad i\,\alpha_2\,\alpha_4 \quad i\,\alpha_3\,\alpha_4 \\
i\,\alpha_2\,\alpha_3 \quad i\,\alpha_3\alpha_1 \quad i\,\alpha_1\,\alpha_2 \quad i\,\alpha_1\,\alpha_2\,\alpha_3 \\
\alpha_1\,\alpha_2\,\alpha_3\,\alpha_4
\end{array} \qquad (22)$$

Dans ce tableau, on a multiplié par i les produits des α_i quand ils sont antihermitiques de façon à obtenir des opérateurs hermitiques. On pourrait naturellement encore en

permutant les α_i obtenir d'autres opérateurs tels que $i\alpha_1\alpha_1$ par exemple, mais chacun de ces nouveaux opérateurs serait égal, ou égal et de signe contraire, à l'un de ceux qui figurent dans le tableau.

En multipliant par un facteur convenable certains opérateurs du tableau (22), nous trouvons des opérateurs qui correspondent à des grandeurs déjà étudiées. Ainsi, d'après les formules (7) et (8) du chapitre XII, les opérateurs de la deuxième ligne, multipliés les trois premiers par ec, le dernier par $-e$, correspondent aux composantes du courant électrique moyen et à la densité électrique moyenne.

Les opérateurs de la troisième ligne, multipliés par le magnéton de Bohr $\frac{eh}{4\pi m_0 c}$, correspondent d'après les equations (9) du chapitre XIV aux coordonnées du moment magnétique propre et du moment électrique propre de l'électron. Enfin les trois premiers opérateurs de la quatrième ligne, multipliés par $\frac{h}{4\pi}$, correspondent suivant les formules (42) du chapitre XV aux trois composantes S_x, S_y et S_z du moment de rotation propre. D'après ce qui a été dit à la fin du dernier chapitre, nous pressentons aussi que l'opérateur α_4 multiplié par $-m_0$ pourrait bien correspondre à la masse propre.

Avec les 16 opérateurs du tableau (22), nous pouvons former 16 densités de valeur moyenne ayant un caractère tensoriel et toutes réelles que nous écrirons comme il suit :

$$\left.\begin{array}{ll}
\Omega_1 = \sum_1^4 {}_k \Psi_k^* \alpha_4 \Psi_k & (a) \\[2ex]
\left.\begin{array}{ll}
i_1 = ec \sum_1^4 {}_k \Psi_k^* \alpha_1 \Psi_k & i_2 = ec \sum_1^4 {}_k \Psi_k^* \alpha_2 \Psi_k \\
i_3 = ec \sum_1^4 {}_k \Psi_k^* \alpha_3 \Psi_k & i_4 = - ec \sum_1^4 {}_k \Psi_k^* \, 1 \, \Psi_k
\end{array}\right\} & (b) \\[2ex]
\left.\begin{array}{ll}
\mu_{32} = \dfrac{eh}{4\pi m_0 c}\, i \sum_1^4 {}_k \Psi_k^* \alpha_2 \alpha_3 \alpha_4 \Psi_k & \mu_{14} = \dfrac{eh}{4\pi m_0 c}\, i \sum_1^4 {}_k \Psi_k^* \alpha_1 \alpha_4 \Psi_k \\
\mu_{13} = \dfrac{eh}{4\pi m_0 c}\, i \sum_1^4 {}_k \Psi_k^* \alpha_3 \alpha_1 \alpha_4 \Psi_k & \mu_{24} = \dfrac{eh}{4\pi m_0 c}\, i \sum_1^4 {}_k \Psi_k^* \alpha_2 \alpha_4 \Psi_k \\
\mu_{21} = \dfrac{eh}{4\pi m_0 c}\, i \sum_1^4 {}_k \Psi_k^* \alpha_1 \alpha_2 \alpha_4 \Psi_k & \mu_{34} = \dfrac{eh}{4\pi m_0 c}\, i \sum_1^4 {}_k \Psi_k^* \alpha_3 \alpha_4 \Psi_k
\end{array}\right\} & (c) \\[2ex]
\left.\begin{array}{ll}
\sigma_1 = \dfrac{h}{4\pi}\, i \sum_1^4 {}_k \Psi_k^* \alpha_2 \alpha_3 \Psi_k ; & \sigma_2 = \dfrac{h}{4\pi}\, i \sum_1^4 {}_k \Psi_k^* \alpha_3 \alpha_1 \Psi_k \\
\sigma_3 = \dfrac{h}{4\pi}\, i \sum_1^4 {}_k \Psi_k^* \alpha_1 \alpha_2 \Psi_k ; & \sigma_4 = \dfrac{h}{4\pi}\, i \sum_1^4 {}_k \Psi_k^* \alpha_1 \alpha_2 \alpha_3 \Psi_k
\end{array}\right\} & (d) \\[2ex]
\Omega_2 = \sum_1^4 {}_k \Psi_k^* \alpha_1 \alpha_2 \alpha_3 \alpha_4 \Psi_k & (e)
\end{array}\right\} \quad (23)$$

Les grandeurs Ω_1 et Ω_2 sont invariantes pour les transformations de Lorentz : nous les connaissions déjà. Envisagées comme densités de valeurs moyennes, ce sont des densités invariantes. La signification physique de Ω_2, si elle existe, est encore inconnue. Celle de la grandeur Ω_1 sera discutée plus loin.

Les quatre quantités (b) forment un quadrivecteur

d'espace-temps [1]. Nous savons déjà qu'elles correspondent aux quatre composantes du « Courant électrique d'Univers » bien connu en Relativité. En intégrant la composante i_4 dans tout l'espace, on obtient une grandeur invariante qui est, au facteur c près, la charge électrique totale $-e$ de l'électron.

Les six quantités (c) sont les six composantes distinctes d'un tenseur antisymétrique du second ordre. Nous les connaissons bien : ce sont les densités de moment magnétique et de moment électrique étudiées au chapitre XIV, paragr. 3.

Les quatre quantités (d) se transforment comme les composantes d'un vecteur d'espace-temps. Les trois premières σ_x, σ_y, σ_z sont, nous le savons déjà les densités de valeur moyenne du moment de rotation propre $\vec{S}$ (spin). La composante de temps σ_4 complète le vecteur d'espace-temps : son interprétation physique n'apparaît pas très clairement.

5. ***Remarques sur le vecteur*** $\vec{\sigma}$. — Le vecteur d'espace-temps $\vec{\sigma}$ peut, d'après ce qui précède, être nommé le « vecteur spin ». La longueur $|\sigma|$ d'un tel vecteur d'espace-temps est donnée par définition par la formule :

$$|\sigma|^2 = \sigma_4^2 - \sigma_1^2 - \sigma_2^2 - \sigma_3^2 \tag{24}$$

Si l'on fait le calcul, le résultat en est :

$$|\sigma|^2 = -(\Omega_1^2 + \Omega_2^2)\frac{h^2}{16\,\pi^2}. \tag{25}$$

Le spin est donc un vecteur du genre « espace » comme on dit en théorie de Relativité. Ceci le différencie nettement du vecteur Courant qui, d'après la formule (14) du chapitre XII, est du genre temps.

Reprenons maintenant l'onde plane monochromatique avec les fonctions d'onde :

---

[1] Dans le tableau (23), les composantes d'indice 4 se rapportent à la quatrième variable d'espace-temps, soit $x_4 = ct$. On a donc $i_4 = c\rho$.

$$
\begin{aligned}
\Psi_1 &= -\frac{A p_z}{\frac{W}{c} + m_0 c} \, e^{\frac{2\pi i}{h}(Wt - p_z z)} \\
\Psi_2 &= \frac{B p_z}{\frac{W}{c} + m_0 c} \, e^{\frac{2\pi i}{h}(Wt - p_z z)} \\
\Psi_3 &= A \, e^{\frac{2\pi i}{h}(Wt - p_z z)} \\
\Psi_4 &= B \, e^{\frac{2\pi i}{h}(Wt - p_z z)}
\end{aligned}
\tag{26}
$$

avec $W = c\sqrt{m_0^2 c^2 + p_z^2}$.

Faisons le calcul explicite des composantes de $\overrightarrow{\sigma}$ dans ce cas particulier ; nous trouvons :

$$
\begin{aligned}
\sigma_1 &= \frac{h}{4\pi}(A^*B + AB^*)\left(1 - \frac{p_z^2}{\left(\frac{W}{c} + m_0 c\right)^2}\right) \\
\sigma_2 &= \frac{h}{4\pi} i(A^*B - B^*A)\left(1 - \frac{p_z^2}{\left(\frac{W}{c} + m_0 c\right)^2}\right) \\
\sigma_3 &= \frac{h}{4\pi}(AA^* - BB^*)\left(1 + \frac{p_z^2}{\left(\frac{W}{c} + m_0 c\right)^2}\right) \\
\sigma_4 &= \frac{h}{4\pi} 2\,(AA^* - BB^*) \cdot \frac{p_z}{\frac{W}{c} + m_0 c}
\end{aligned}
\tag{27}
$$

Comme les composantes du quadrivecteur « Courant électrique d'Univers » sont pour l'onde plane :

$$
i_1 = j_x = 0 \qquad i_2 = j_y = 0 \qquad i_3 = j_z = -e\rho \frac{c^2 p_z}{W}
$$
$$
i_4 = c\delta = -e\rho c
$$

comme cela résulte des formules du chapitre XII, parag. 4, nous trouvons aisément :

$$
\left(\overrightarrow{i} \cdot \overrightarrow{\sigma}\right) = i_4\sigma_4 - i_1\sigma_1 - i_2\sigma_2 - i_3\sigma_3 = 0. \tag{28}
$$

Ainsi le produit scalaire dans l'espace-temps des vecteurs $\vec{i}$ et $\vec{\sigma}$ se trouve ici nul. Donc pour l'onde plane monochromatique, les deux quadrivecteurs « Courant d'Univers » et « Spin » sont orthogonaux dans l'espace-temps (ce qui ne veut pas dire, bien entendu, que les vecteurs d'espace correspondant soient perpendiculaires).

On peut aussi écrire (28) sous la forme :

$$\sigma_4 = \sigma_3 \frac{p_z c}{W} = \sigma_1 v_1 + \sigma_2 v_2 + \sigma_3 v_3 . \qquad (28bis)$$

La composante σ_4 s'interprète donc ici comme le produit scalaire *dans l'espace* des vecteurs « densité de moment de rotation propre » et « vitesse ».

Pour une vitesse voisine de c, le rapport $\frac{p_z}{\frac{W}{c} + m_0 c}$ est voisin de 1 et, d'après (27), les composantes σ_x et σ_y sont presque nulles. En rapprochant ce résultat de ce qui a été dit à la fin du dernier chapitre, on voit que dans la Mécanique de Dirac, une particule animée d'une vitesse voisine de c et associée à une onde plane monochromatique a ses trois vecteurs d'espace $\vec{I}$, $\vec{J}$ et $\vec{\sigma}$ mutuellement perpendiculaires, le dernier étant dirigé suivant la normale à l'onde [1].

---

[1] On pourrait être surpris de voir que le moment de rotation propre $\vec{S}$ de l'électron de Dirac ne coïncide pas toujours en direction avec le moment magnétique $\mathfrak{M}$ et croire que cela est contradictoire avec la relation :

$$\frac{\mathfrak{M}}{S} = -\frac{e}{m_0 c} \qquad (*)$$

par laquelle s'exprime le double magnétisme de l'électron. Mais on doit remarquer que la relation (*) n'est valable que pour les systèmes de référence où l'électron est en repos car $\vec{\mathfrak{M}}$ et $\vec{S}$ ne se transforment pas de même pour une transformation de Lorentz. En effet pour un moment magnétique, ce sont les *densités* des composantes qui se transforment comme les composantes 23, 31 et 12 d'un tenseur ; pour un moment de rotation au contraire, ce sont les composantes elles-mêmes qui se transforment de cette façon.

Nous terminerons ces remarques sur le spin en signalant une relation générale entre les composantes de ce vecteur et l'invariant Ω_2. Cette relation due à MM. Uhlenbeck et Laporte s'écrit avec nos notations :

$$\frac{1}{c}\frac{\partial \sigma_4}{\partial t}+\frac{\partial \sigma_1}{\partial x}+\frac{\partial \sigma_2}{\partial y}+\frac{\partial \sigma_3}{\partial z}=\sum_1^4{}_i \frac{\partial \sigma_i}{\partial x_i}=-m_0 c\,\Omega_2\,. \tag{29}$$

Dans le cas de l'onde plane, elle est identiquement satisfaite parce que tous les termes en sont séparément nuls. On démontre aisément la relation (28) en partant de l'équation de Dirac et de sa conjuguée ; son sens physique véritable reste inconnu.

6. ***Remarques sur l'invariant*** Ω_1 ***et l'opérateur*** $-\mathbf{m}_0\alpha_4$. — Nous avons déjà vu qu'à la grandeur physique « masse propre » correspondait en un certain sens dans les équations de Dirac l'opérateur $-m_0\alpha_4$. Si nous admettons cette correspondance, la densité de valeur moyenne qu'on en déduit est :

$$-m_0\sum_1^4{}_k \Psi_k^*\alpha_4\Psi_k=-m_0\,\Omega_1\,. \tag{30}$$

Elle est invariante. Si l'on intègre cette densité dans l'espace, on obtient la valeur moyenne

$$\overline{m_0}=-m_0\int\!\!\int\!\!\int_{+\infty}^{-\infty}\sum_1^4{}_k \Psi_k^*\alpha_4\Psi\,d\tau=-m_0\int\!\!\int\!\!\int_{-\infty}^{-\infty}\Omega_1\,d\tau \tag{31}$$

ou en explicitant Ω_1 :

$$\overline{m_0}=m_0\int\!\!\int\!\!\int_{-\infty}^{+\infty}(\Psi_3^*\Psi_3+\Psi_4^*\Psi_4-\Psi_1^*\Psi_1-\Psi_2^*\Psi_2)\,d\tau \tag{32}$$

formule qui confirme bien l'idée que les ondes d'indices 3 et 4 correspondent à la masse propre $+m_0$ et les ondes d'indices 1 et 2 à la masse propre $-m_0$.

La valeur moyenne (31) n'est pas invariante ; si l'on veut

en déduire une quantité invariante, on devra intégrer par rapport à la quatrième variable d'espace-temps $x_4 = ct$; on obtient ainsi :

$$\int^t \overline{m_0}\, c\, dt = -m_0 c \int^t dt \int\int\int_{-\infty}^{+\infty} \Omega_1\, dx\, dy\, dz \tag{33}$$

Pour tâcher de pénétrer un peu la signification physique de ces formules, reprenons encore le cas de l'onde plane monochromatique (26) et calculons la valeur qu'a $\overline{m_0}$ en ce cas. Nous trouvons aisément d'après (30) :

$$\begin{aligned} -m_0 \Omega_1 &= m_0 (AA^* + BB^*) \left[1 - \frac{p_z^2}{\left(\frac{W}{c} + m_0 c\right)^2}\right] \\ &= m_0 (AA^* + BB^*) \frac{2 m_0 c^2}{W + m_0 c^2} \end{aligned} \tag{34}$$

Pour avoir $\overline{m_0}$, il faut intégrer ce qui donne en tenant compte de la normalisation des Ψ (¹) :

$$\overline{m_0} = m_0 \cdot \frac{2 m_0 c^2}{W + m_0 c^2} \cdot \frac{1}{1 + \frac{p_z^2}{\left(\frac{W}{c} + m_0 c\right)^2}} \tag{35}$$

ou encore :

$$\overline{m_0} = m_0 \cdot \frac{m_0 c^2}{W} = m_0 \sqrt{1 - \beta^2} \tag{36}$$

βc étant la vitesse qui correspond suivant la Dynamique relativiste à l'énergie W.

L'intégrale invariante (33) n'est donc pas autre chose (à une constante près) que l'*intégrale d'action* :

$$\int^t m_0 c \sqrt{1 - \beta_2\, dt} \tag{37}$$

du point matériel dans la Dynamique de Relativité. Nous avons donc le théorème suivant : « Pour un électron de Dirac associé à une onde monochromatique plane, l'intégrale dans le temps de la valeur moyenne de la masse propre coïncide avec l'intégrale d'action de la Dynamique d'Einstein. »

(¹) Voir la note page 184.

Grandeurs physiques		Opérateurs	Densités de valeur moyenne	Variance relativiste de la densité
Masse propre ?		$(-m_0)\,\alpha_4$	$-m_0 \sum_{1}^{4}{}_k \Psi_k^* \,\alpha_4\, \Psi_k = -m_0\,\Omega_1$	Invariant
Charge électrique		$-e\,.\,1$	$\delta = -\rho e = -e \sum_{1}^{4}{}_k \Psi_k^*\,.\,1\,\Psi_k$	Vecteur d'espace-temps (Courant d'Univers)
Courant électrique		$e\,c\,\alpha_1$	$j_x = -\rho\, e\, u_x = e\, c \sum_{1}^{4}{}_k \Psi_k^*\,.\,\alpha_1\,\Psi_k$	
		$e\,c\,\alpha_2$	$j_y = -\rho\, e\, u_y = e\, c \sum_{1}^{4}{}_k \Psi_k^*\,\alpha_2\,\Psi_k$	
		$e\,c\,\alpha_3$	$j_z = -\rho\, e\, u_z = e\, c \sum_{1}^{4}{}_k \Psi_k^*\,\alpha_3\,\Psi_k$	
Moment magnétique	M_x	$\frac{e\,h}{4\pi m_0 c}\,.\,i\,\alpha_2\,\alpha_3\,\alpha_4$	$I_x = \frac{e\,h}{4\pi m_0 c}\, i \sum_{1}^{4}{}_k \Psi_k^*\,\alpha_2\,\alpha_3\,\alpha_4\,\Psi_k$	Tenseur antisymétrique du second rang 1re partie
	M_y	,, $.\,i\,\alpha_3\,\alpha_1\,\alpha_4$	$I_y = \frac{e\,h}{4\pi m_0 c}\, i \sum_{1}^{4}{}_k \Psi_k^*\,\alpha_3\,\alpha_1\,\alpha_4\,\Psi_k$	
	M_z	,, $.\,i\,\alpha_1\,\alpha_2\,\alpha_4$	$I_z = \frac{e\,h}{4\pi m_0 c}\, i \sum_{1}^{4}{}_k \Psi_k^*\,\alpha_1\,\alpha_2\,\alpha_4\,\Psi_k$	

Tableau des grandeurs et densités associées à l'électron.

Grandeurs physiques		Opérateurs	Densités de valeur moyenne	Variance relativiste de la densité
Moment électrique	P_x	$\frac{e\,h}{4\,\pi\,m_0\,c} \cdot i\,\alpha_1\,\alpha_4$	$J_x = \frac{e\,h}{4\,\pi\,m_0\,c}\, i \sum_{1}^{4}{}_k \Psi'^*_k\, \alpha_1\, \alpha_4\, \Psi'_k$	Tenseur antisymétrique du second rang 2^me partie
	P_y	,, $\cdot\, i\,\alpha_2\,\alpha$	$J_y = \frac{e\,h}{4\,\pi\,m_0\,c}\, i \sum_{1}^{4}{}_k \Psi'^*_k\, \alpha_2\, \alpha_4\, \Psi'_k$	
	P_z	,, $\cdot\, i\,\alpha_3\,\alpha_4$	$J_z = \frac{e\,h}{4\,\pi\,m_0\,c}\, i \sum_{1}^{4}{}_k \Psi'^*_k\, \alpha_3\, \alpha_4\, \Psi'_k$	
Moment de rotation propre (spin)	S_x	$\frac{h}{4\,\pi}\, i\,\alpha_2\,\alpha_3$	$\sigma_x = \frac{h}{4\,\pi}\, i \sum_{1}^{4}{}_k \Psi'^*_k\, \alpha_2\, \alpha_3\, \Psi'_k$	Tenseur complètement antisymétrique du 3^e rang = Vecteur d'espace temps
	S_y	,, $i\,\alpha_3\,\alpha_1$	$\sigma_y = \frac{h}{4\,\pi}\, i \sum_{1}^{4}{}_k \Psi'^*_k\, \alpha_3\, \alpha_1\, \Psi'_k$	
	S_z	,, $i\,\alpha_1\,\alpha_3$	$\sigma_z = \frac{h}{4\,\pi}\, i \sum_{1}^{4}{}_k \Psi'^*_k\, \alpha_1\, \alpha_2\, \Psi'_k$	
?	S_4	,, $i\,\alpha_1\,\alpha_2\,\alpha_3$	$\sigma_4 = \frac{h}{4\,\pi}\, i \sum_{1}^{4}{}_k \Psi'^*_k\, \alpha_1\, \alpha_2\, \alpha_3\, \Psi'_k$	
?	?	$\alpha_1\,\alpha_2\,\alpha_3\,\alpha_4$	$\Omega_2 = \sum_{1}^{4}{}_k \Psi'^*_k\, \alpha_1\, \alpha_2\, \alpha_3\, \alpha_4\, \Psi'_k$	Tenseur complètement antisymétrique du 4^e rang = Invariant

Tableau des grandeurs et densités associées à l'électron.

$$\alpha_4 = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{vmatrix}$$

$$\alpha_1 = \begin{vmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{vmatrix} ; \quad \alpha_2 = \begin{vmatrix} 0 & 0 & 0 & i \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{vmatrix} ; \quad \alpha_3 = \begin{vmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{vmatrix} ; \quad 1 = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}$$

$$i\,\alpha_2\,\alpha_3\,\alpha_4 = \begin{vmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{vmatrix} ; \quad i\,\alpha_3\,\alpha_1\,\alpha_4 = \begin{vmatrix} 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{vmatrix} ; \quad i\,\alpha_1\,\alpha_2\,\alpha_4 = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}$$

$$i\,\alpha_1\,\alpha_4 = \begin{vmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \end{vmatrix} ; \quad i\,\alpha_2\,\alpha_4 = \begin{vmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{vmatrix} ; \quad i\,\alpha_3\,\alpha_4 = \begin{vmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & i \\ i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{vmatrix}$$

$$i\,\alpha_2\,\alpha_3 = \begin{vmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{vmatrix} ; \quad i\,\alpha_3\,\alpha_1 = \begin{vmatrix} 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & -i & 0 \end{vmatrix} ; \quad i\,\alpha_1\,\alpha_2 = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{vmatrix} ; \quad i\,\alpha_1\,\alpha_2\,\alpha_3 = \begin{vmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{vmatrix}$$

$$\alpha_1\,\alpha_2\,\alpha_3\,\alpha_4 = \begin{vmatrix} 0 & 0 & i & 0 \\ 0 & 0 & 0 & i \\ -i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{vmatrix}$$

Tableau des 16 opérateurs hermitiques fondamentaux de la théorie de Dirac

TROISIÈME PARTIE

APPLICATIONS DE LA THÉORIE DE DIRAC
CRITIQUES ET COMPLÉMENTS DIVERS

CHAPITRE XVII

Explication de la structure fine par la théorie de Dirac

1. *Equations d'ondes pour le mouvement d'un électron dans un champ central* [1]

NOUS nous proposons de montrer dans ce chapitre que la théorie de Dirac fournit une bonne explication de la structure fine des spectres optiques et des spectres X sans soulever les mêmes difficultés que l'ancienne théorie de structure fine de Sommerfeld.

Considérons un électron qui se meut dans un champ statique central dérivant d'un potentiel $V(r)$. Les équations de Dirac pour cet électron seront :

$$a)\quad \frac{2\pi i}{h}\left[\frac{W+eV}{c}+m_0 c\right]\Psi_1-\left(\frac{\partial}{\partial x}+i\frac{\partial}{\partial y}\right)\Psi_4-\frac{\partial}{\partial z}\Psi_3=0$$

$$b)\quad \frac{2\pi i}{h}\left[\frac{W+eV}{c}+m_0 c\right]\Psi_2-\left(\frac{\partial}{\partial x}-i\frac{\partial}{\partial y}\right)\Psi_3+\frac{\partial}{\partial z}\Psi_4=0$$

$$c)\quad \frac{2\pi i}{h}\left[\frac{W+eV}{c}-m_0 c\right]\Psi_3-\left(\frac{\partial}{\partial x}+i\frac{\partial}{\partial y}\right)\Psi_2-\frac{\partial}{\partial z}\Psi_1=0$$

$$d)\quad \frac{2\pi i}{h}\left[\frac{W+eV}{c}-m_0 c\right]\Psi_4-\left(\frac{\partial}{\partial x}-i\frac{\partial}{\partial y}\right)\Psi_1+\frac{\partial}{\partial z}\Psi_2=0 \qquad (1)$$

[1] Nous suivons ici la méthode employée par M. C. G. Darwin (*Proc. Roy. Soc. A.*, vol. 118, 1928, p. 554).

Il est naturel de chercher à exprimer chacun des Ψ_k comme un produit d'une fonction sphérique de Laplace par une fonction du rayon vecteur.

Rappelons que les fonctions sphériques de Laplace ont la forme suivante :

$$Y_l^m(\theta,\varphi) = C\, e^{\pm i\varphi m}\, P_l^m(\cos\theta) = C\, e^{\pm i m\varphi} \sin^m\theta \frac{d^{l+m}}{(d\cos\theta)^{l+m}}(1-\cos^2\theta)^l \qquad (2)$$

La constante C étant arbitraire, choisissons-la, avec M. Darwin, de façon à avoir :

$$Y_l^m(\theta,\varphi) = (l-m)!\, e^{i m\varphi} \frac{d^{l+m}}{(d\cos\theta)^{l+m}} \left(\frac{(1-\cos^2\theta)^l}{2^l \cdot l\,!}\right) \cdot \sin^m\theta \quad (3)$$

avec $l = 0, 1, 2 \ldots$ et $m = -l, -(l-1) \ldots +l$. Remarquons d'ailleurs que la fonction (3) n'est pas normée sur la surface de la sphère de rayon un.

En se servant des formules de transformation des coordonnées rectangulaires en coordonnées polaires, on démontre les relations suivantes où $f(r)$ désigne une fonction quelconque du rayon vecteur :

$$\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right) f\, Y_l^m = \frac{1}{2l+1}\left\{\left[\frac{df}{dr} - \frac{l}{r}f\right] Y_{l+1}^{m+1} - (l-m)(l-m-1)\left(\frac{df}{dr} + \frac{l+1}{r}f\right) Y_{l-1}^{m+1}\right\}$$

$$\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right) f\, Y_l^m = \frac{1}{2l+1}\left\{-\left[\frac{df}{dr} - \frac{l}{r}f\right] Y_{l+1}^{m-1} + (l+m)(l+m-1)\left(\frac{df}{dr} + \frac{l+1}{r}f\right) Y_{l-1}^{m-1}\right\} \qquad (4)$$

$$\frac{\partial}{\partial z} f\, Y_l^m = \frac{1}{2l+1}\left\{\left(\frac{df}{dr} - \frac{l}{r}f\right) Y_{l+1}^{m} + (l+m)(l-m)\left(\frac{df}{dr} + \frac{l+1}{r}f\right) Y_{l-1}^{m}\right\}$$

A l'aide des formules (4), nous allons pouvoir expliciter les dérivées qui figurent dans (1) puisque nous supposons que chaque Ψ_k est le produit d'une fonction de r par une

fonction de Laplace. Faisons alors l'hypothèse que la fonction Ψ_3 est proportionnelle à la fonction Y_l^m avec une valeur donnée de l et de m.

En regardant maintenant l'équation (c) de (1), nous voyons que les termes en Y_l^m provenant de $-\frac{\partial \Psi_1}{\partial z}$ et de $-\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)\Psi_2$ doivent détruire le terme en Ψ_3 et que les autres termes doivent se détruire entre eux. Il en résulte que Ψ_1 et Ψ_2 doivent dépendre de la même fonction du rayon vecteur et être proportionnelles respectivement soit à Y_{l+1}^{m} et Y_{l+1}^{m-1}, soit à Y_{l-1}^{m} et Y_{l-1}^{m-1}.

De même l'équation (a) de (1) montre que Ψ_3 et Ψ_4 doivent dépendre d'une même fonction du rayon vecteur et que Ψ_4 doit être proportionnelle à Y_{l}^{m-1}.

Finalement nous sommes amenés à envisager d'abord une solution que nous pouvons écrire [1]:

$$\left.\begin{aligned} \Psi_1 &= ia_1\,F_+(r)\,Y_{l+1}^{m}(\theta,\varphi)\ ; & \Psi_2 &= ia_2\,F_+(r)\,Y_{l+1}^{m-1}(\theta,\varphi) \\ \Psi_3 &= a_3\,G_+(r)\,Y_{l}^{m}(\theta,\varphi) & \Psi_4 &= a_4\,G_+(r)\,Y_{l}^{m-1}(\theta,\varphi) \end{aligned}\right\} \quad \text{(I)}$$

Introduisons les formes (I) dans les équations (a) et (b) de (1). Nous obtenons :

$$-\frac{2\pi}{h}\left[\frac{W+eV}{c}+m_0c\right]a_1\,F_+\,Y_{l+1}^{m} - \frac{a_4}{2\,l+1}$$
$$\left\{\left(\frac{dG_+}{dr}-\frac{l}{r}G_+\right)Y_{l+1}^{m} - (l-m)\,(l-m+1)\right.$$
$$\left.\left(\frac{dG_+}{dr}+\frac{l+1}{r}G_+\right)Y_{l-1}^{m}\right\} - \frac{a_3}{2\,l+1}\left\{\left(\frac{dG_+}{dr}-\frac{l}{r}G_+\right)Y_{l+1}^{m}\right.$$
$$\left. + (l+m)\,(l-m)\left(\frac{dG_+}{dr}+\frac{l+1}{r}G_+\right)Y_{l-1}^{m}\right\} = 0$$

(5)

[1] Nous introduisons le facteur i dans Ψ_1 et Ψ_2 pour que a_1 et a_2 soient réels. Cf. formules (9).

$$-\frac{2\pi}{h}\left[\frac{W+eV}{c}+m_0c\right]a_2\,F_+\,Y_{l+1}^{m-1}-\frac{a_3}{2\,l+1}$$

$$\Big\{-\left(\frac{dG_+}{dr}-\frac{l}{r}\,G_+\right)Y_{l+1}^{m-1}+(l+m)\,(l+m-1)$$

$$\left(\frac{dG_+}{dr}+\frac{l+1}{r}\,G_+\right)Y_{l-1}^{m-1}\Big\}+\frac{a_4}{2\,l+1}\Big\{\left(\frac{dG_+}{dr}-\frac{l}{r}\,G_+\right)Y_{l+1}^{m-1}$$

$$+(l+m-1)\,(l-m+1)\left(\frac{dG_+}{dr}+\frac{l+1}{r}G_+\right)Y_{l-1}^{m-1}\Big\}=0.$$

Pour faire disparaître les termes en Y_{l-1}^{m} dans la première équation (5) et les termes en Y_{l-1}^{m-1} dans la seconde, il suffit de poser :

$$\frac{a_3}{a_1}=\frac{l-m+1}{l+m}. \tag{6}$$

D'autre part, substituons les formes (I) dans les équations (c) et (d) de (1) ; il vient :

$$\frac{2\pi i}{h}\left[\frac{W+eV}{c}-m_0c\right]a_3\,G_+\,Y_{l}^{m}-\frac{ia_2}{2\,l+3}$$

$$\Big\{\left(\frac{dF_+}{dr}-\frac{l+1}{r}\,F_+\right)Y_{l+2}^{m}\,(l-m+2)\,(l-m+1)$$

$$\left(\frac{dF_+}{dr}+\frac{l+2}{r}\,F_+\right)Y_{l}^{m}\Big\}-\frac{ia_1}{2\,l+3}\Big\{\left(\frac{dF_+}{dr}-\frac{l+1}{r}\,F_+\right)Y_{l+2}^{m}$$

$$+(l+m+1)\,(l-m+1)\left(\frac{dF_+}{dr}+\frac{l+2}{r}\,F_+\right)Y_{l}^{m}\Big\}=0$$

(7)

$$\frac{2\pi i}{h}\left[\frac{W+eV}{c}-m_0c\right]a_4\,G_+\,Y_{l}^{m-1}-\frac{ia_1}{2\,l+3}\Big\{-\Big(\frac{dF_+}{dr}-$$

$$-\frac{l+1}{r}\,F_+\Big)\,Y_{l+2}^{m-1}+(l+m+1)\,(l+m)\Big(\frac{dF_+}{dr}+$$

$$+\frac{l+2}{r}\,F_+\Big)\,Y_{l}^{m-1}\Big\}+\frac{ia_2}{2\,l+3}\Big\{\left(\frac{dF_+}{dr}-\frac{l+1}{r}\,F_+\right)Y_{l+2}^{m-1}+$$

$$+(l+m)\,(l-m+2)\left(\frac{dF_+}{dr}+\frac{l+2}{r}\,F_+\right)Y_{l}^{m-1}\Big\}=0.$$

Pour faire disparaître les termes en Y_{l+2}^{m} de la première

équation (7) et les termes en Y_{l+2}^{m-1} de la seconde, il faut poser :

$$a_2 = -a_1 \tag{8}$$

Le rapport $\frac{a_1}{a_3}$ est arbitraire car nous pouvons l'incorporer dans $\frac{F_+}{G_+}$. Nous pouvons donc satisfaire aux conditions (6) et (8) en posant :

$$a_1 = 1 \quad a_2 = -1 \quad a_3 = l - m + 1 \quad a_4 = l + m \tag{9}$$

de sorte que $a_3 + a_4 = 2l + 1$.

Dans ces conditions, les deux équations (5) se réduisent à l'équation unique :

$$\frac{2\pi}{h}\left[\frac{W + eV}{c} + m_0 c\right] F_+ + \frac{dG_+}{dr} - \frac{l}{r} G_+ = 0. \tag{10}$$

Les équations (7) se réduisent de même à l'équation unique :

$$-\frac{2\pi}{h}\left[\frac{W + eV}{c} - m_0 c\right] G_+ + \frac{dF_+}{dr} + \frac{l+2}{r} F_+ = 0. \tag{11}$$

La solution (I) est caractérisée par le fait que les indices inférieurs de Y sont l et $l+1$. La moyenne de ces deux valeurs est $l + \frac{1}{2}$ que nous nommerons j. La solution (I) est donc caractérisée par les nombres quantiques l, $j = l + \frac{1}{2}$ et m.

Nous venons de trouver une première solution dans laquelle Ψ_3 est proportionnelle à Ψ_l^m. Mais nous avons vu qu'il peut en exister une autre dans laquelle Ψ_1 est proportionnelle à Y_{l-1}^{m} et Ψ_2 à Y_{l-1}^{m-1}.

Nous sommes donc conduits à envisager une solution (II) de la forme :

$$\begin{aligned} &\Psi_1 = ia_1' F_-(r) Y_{l-1}^{m}(\theta,\varphi); \quad \Psi_2 = ia_2' F_-(r) Y_{l-1}^{m-1}(\theta,\varphi) \\ &\Psi_3 = a_3' G_-(r) Y_{l}^{m}(\theta,\varphi) \qquad \Psi_4 = a_4' G_-(r) Y_{l}^{m-1}(\theta,\varphi) \end{aligned} \tag{II}$$

En substituant les formes (II) dans les équations (1) et en raisonnant comme plus haut, on est amené aux conditions :

$$\frac{a_2'}{a_1'} = -\frac{l+m-1}{l-m} \qquad a_3' = -a_4' \tag{12}$$

analogues à (6) et (8). On satisfera à ces conditions en posant :

$$a'_1 = -(l-m) \quad a_2' = l+m-1 \quad a_3' = 1 \quad a_4' = -1. \tag{13}$$

On trouve alors pour F_- et G_- les équations :

$$\begin{aligned} \frac{2\pi}{h}\left[\frac{W+eV}{c}+m_0c\right]F_- + \frac{dG}{dr} + \frac{l+1}{r}G_- = 0 \\ -\frac{2\pi}{h}\left[\frac{W+eV}{c}-m_0c\right]G_- + \frac{dF_-}{dr} - \frac{l-1}{r}F_- = 0 \end{aligned} \tag{14}$$

analogues à (10) et (11).

Les fonctions d'onde (II) constituent la solution correspondant aux nombres quantiques : $l, j = l - \frac{1}{2}$ et m.

Nous remarquerons dès maintenant que pour $l = 0$ (termes s), la solution (II) n'existe pas car il n'y a pas de fonctions Y^{m}_{-1} à indice négatif.

Nous verrons bientôt que la détermination des fonctions F et G fait intervenir un nombre quantique n, le nombre quantique total, de sorte que toute solution complète est caractérisée par 4 nombres quantiques : n, l, $j = l \pm \frac{1}{2}$, m.

Nous résumerons donc les résultats obtenus dans ce paragraphe par le tableau suivant :

Solution (I) $n, l, j = l + \frac{1}{2}, m$

$$\Psi_1 = i\,F_+\,Y^{m}_{l+1}\,; \quad \Psi_2 = -\,i\,F_+\,Y^{m-1}_{l+1}\,;$$

$$\Psi_3 = (l-m+1)\,G_+\,Y^{m}_{l}\,; \quad \Psi_4 = (l+m)\,G_+\,Y^{m-1}_{l}$$

$$\frac{2\pi}{h}\left[\frac{W+eV}{c}+m_0c\right]F_+ + \frac{dG_+}{dr} - \frac{l}{r}G_+ = 0\,;$$

$$-\frac{2\pi}{h}\left[\frac{W+eV}{c}-m_0c\right]G_+ + \frac{dF_+}{dr} + \frac{l+2}{r}F_+ = 0.$$

Solution (II) $\qquad n,\, l,\, j = l - \frac{1}{2},\, m$

$$\Psi_1 = -i\,(l-m)\,F_-\,Y_{l-1}^{m}\,; \qquad \Psi_2 = i\,(l+m-1)\,F_-\,Y_{l-1}^{m-1}$$

$$\Psi_3 = G_-\,Y_l^{m}\,; \quad \Psi_4 = -\,G_-\,Y_l^{m-1}$$

$$\frac{2\pi}{h}\left[\frac{W+eV}{c}+m_0c\right]F_- + \frac{dG_-}{dr} + \frac{l+1}{r}G_- = 0\,;$$

$$-\frac{2\pi}{h}\left[\frac{W+eV}{c}-m_0c\right]G_- + \frac{dF_-}{dr} - \frac{l-1}{r}F_- = 0.$$

On remarquera sur ce tableau que les équations satisfaites par F_- et G_- se déduisent de celles satisfaites par F_+ et G_+ en changeant l en $-(l+1)$.

2. *Nombre de solutions correspondant à une valeur donnée de* n *et de* l*; moment de rotation de l'électron dans les états stationnaires.* — Il est très intéressant de se demander combien il peut exister d'états stationnaires correspondant à une valeur donnée des nombres quantiques n et l. Il est évident qu'il y en a *a priori* autant que de valeurs permises par m quand n et l sont fixés.

Considérons d'abord les solutions du type (I) et comptons les valeurs possibles de m en remarquant que l'indice supérieur des fonctions Y doit être au plus égal, en valeur absolue, à l'indice inférieur. Il n'y a aucune raison qui s'oppose à prendre $m=0,\ 1,\ \ldots\ l$; il pourrait sembler impossible de prendre $m=l+1$ car alors Ψ_3 contiendrait la fonction inexistante Y_l^{l+1}, mais le coefficient a_3 se trouve nul pour la valeur $m=l+1$ et par suite on peut encore prendre cette valeur de m. Les valeurs de m supérieures à $l+1$ sont par contre inacceptables.

D'autre part on peut prendre sans difficultés les valeurs négatives $-1, -2 \ldots, -(l-1)$. Peut-il prendre la valeur $-l$? Oui car a_4 est alors nul. Mais les valeurs de m inférieures à $-l$ sont inacceptables. Finalement pour les solutions du type (I) telles que $j = l + \frac{1}{2}$, on a $2l+2$ valeurs possibles de m, savoir l valeurs négatives, la valeur zéro, et $l+1$ valeurs positives. Donc il y a en ce cas $2l+2 = 2j+1$ états stationnaires *a priori* distincts.

Prenons les solutions du type (II), m peut prendre les valeurs $0, 1, \ldots, l-1$ et il peut aussi prendre la valeur l qui fait apparaître dans Ψ_1 la fonction inexistante Y_{l-1}^{l} mais qui annule simultanément a_1'. Le nombre m peut aussi prendre les valeurs négatives $-1, -2, \ldots, -(l-2)$ et, de plus, il peut aussi prendre la valeur $-(l-1)$ qui fait apparaître dans Ψ_2 la fonction inexistante Y_{l-1}^{-l} mais annule simultanément a_2'. Les valeurs de m supérieures à l ou inférieures à $-(l-1)$ sont inacceptables. Donc nous avons $2l$ valeurs possibles de m, savoir l valeurs positives, la valeur zéro, et $(l-1)$ valeurs négatives.

Comme pour les solutions du type II on a $j = l - \frac{1}{2}$, il y a ici $2l = 2j+1$ états stationnaires *a priori* distincts.

En résumé pour les solutions (II) comme pour les solutions (I), il y a $(2j+1)$ solutions distinctes. Nous verrons plus loin comment ce résultat permet de justifier la règle de Stoner relative à la distribution des électrons dans l'atome.

Nous allons maintenant chercher à caractériser les types (I) et (II) de solutions par les moments de rotation correspondants. Pour cela, nous partirons du résultat précédemment démontré : l'opérateur intégrale première qui correspond en théorie de Dirac au moment de rotation autour de oz est non pas $M_z = \frac{h}{2\pi i}\left(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\right) = -\frac{h}{2\pi i}\frac{\partial}{\partial \varphi}$, mais bien

$N_z = M_z + S_z = -\frac{h}{2\pi i}\frac{\partial}{\partial \varphi} - \frac{h}{4\pi i}\alpha_1 \alpha_2$. Or, si l'on remarque que $\frac{\partial Y_l^m}{\partial \varphi} = im\, Y_l^m$, on voit aisément que l'on a, aussi bien pour une solution du type (II) que pour une solution du type (I), les formules :

$$N_z(\Psi_k) = -\frac{h}{2\pi i}\frac{\partial \Psi_k}{\partial \varphi} - \frac{h}{4\pi i}\alpha_1 \alpha_2 \Psi_k = -\left(m - \frac{1}{2}\right)\frac{h}{2\pi}\Psi_k \tag{15}$$

pour $k = 1, 2, 3, 4$. On voit donc que pour une solution correspondant au nombre quantique m, qu'elle soit du type (I) ou du type (II), le moment de rotation total N_z est égal à $-\left(m - \frac{1}{2}\right)\frac{h}{2\pi}$. Voyons maintenant quelles sont les valeurs extrêmes que ce moment peut prendre pour une valeur donnée de l. Pour une solution du type (I), le nombre m peut varier de $-l$ à $+(l+1)$ et par suite le moment N_z peut varier de $-\left(l+\frac{1}{2}\right)\frac{h}{2\pi}$ à $+\left(l+\frac{1}{2}\right)\frac{h}{2\pi}$; le type (I) de solution correspond donc à un moment N_z dont la valeur absolue maxima est $\left(l+\frac{1}{2}\right)\frac{h}{2\pi}$. Pour une solution du type (II), le nombre m peut varier, pour l donné, de $-(l-1)$ à $+l$; le moment N_z peut donc varier de

$$-\left(l - \frac{1}{2}\right)\frac{h}{2\pi} \text{ à } + \left(l - \frac{1}{2}\right)\frac{h}{2\pi}$$

et une solution du type (II) se trouve ainsi correspondre à une valeur absolue maxima de N_z égale à $\left(l - \frac{1}{2}\right)\frac{h}{2\pi}$.

Or, pour employer l'ancien langage, on doit dire que $l\frac{h}{2\pi}$ est le moment de rotation « orbital » de l'électron. Nous apercevons alors que les solutions du type (I) peuvent être caractérisées en disant qu'elles correspondent aux cas où spin et moment orbital sont parallèles et de même sens : les

solutions du type (II) correspondent au contraire aux cas où spin et moment orbital sont parallèles et de sens contraire.

3. *Calcul des niveaux d'énergie pour l'atome d'hydrogène.* — Nous allons calculer les niveaux d'énergie pour l'atome d'hydrogène. Dans ce cas il faut poser dans les équations du premier paragraphe $V = \frac{e}{r}$. Nous allons examiner successivement le cas des solutions du type (I), puis celui des solutions du type (II).

a) *Solutions du type (I).*

Pour les solutions du type (I), nous avons les deux équations :

$$\begin{aligned} \frac{2\pi}{h}\left[\frac{W}{c} + \frac{e^2}{cr} + m_0 c\right] F_+ + \frac{dG_+}{dr} - \frac{l}{r} G_+ = 0 \\ -\frac{2\pi}{h}\left[\frac{W}{c} + \frac{e^2}{cr} - m_0 c\right] G_+ + \frac{dF_+}{dr} - \frac{l+2}{r} F_+ = 0. \end{aligned} \tag{16}$$

Posons :

$$A^2 = \frac{2\pi}{h}\left(\frac{W}{c} + m_0 c\right) \qquad B^2 = \frac{2\pi}{h}\left(m_0 c - \frac{W}{c}\right) \tag{17}$$

et introduisons la constante de structure fine $\alpha = \frac{2\pi e^2}{ch}$.

Nous voulons trouver les niveaux du spectre discontinu. Ces niveaux correspondent, nous le savons, à une énergie $E = W - m_0 c^2$ qui est négative. La quantité B^2 est donc positive et B est réelle. Avec ces notations, les équations (16) deviennent :

$$\begin{aligned} \left(A^2 + \frac{\alpha}{r}\right) F_+ + \frac{dG_+}{dr} - \frac{l}{r} G_+ = 0 \\ \left(B^2 - \frac{\alpha}{r}\right) G_+ + \frac{dF_+}{dr} + \frac{(l+2)}{r} F_+ = 0. \end{aligned} \tag{18}$$

Asymptotiquement, pour r très grand, on a :

$$A^2 F_+ + \frac{dG_+}{dr} = 0 \,; \quad B^2 G_+ + \frac{dF_+}{dr} = 0 \tag{19}$$

d'où l'on tire :

$$\frac{d^2\,G_+}{dr^2} = A^2B^2G_+ \qquad \frac{d^2F_+}{dr^2} = A^2B^2F_- \tag{20}$$

et en intégrant :

$$G_+ = e^{\pm ABr} \qquad F_+ = e^{\pm ABr} \tag{21}$$

Nous devons prendre le signe — pour que F_- et G_+ soient nulles à l'infini, d'où :

$$G_{+\,(\text{asymp.})} = C^{te}\,e^{-ABr} \qquad F_{+\,(\text{asymp.})} = C^{te}\,e^{-ABr}. \tag{22}$$

Ces formes asymptotiques nous suggèrent d'essayer de substituer dans (18) les formes suivantes :

$$F_+ = e^{-ABr}\left[a_0 r^{\gamma} + a_1 r^{\gamma+1} \cdots + a_s r^{\gamma+s} + \cdots\right];$$
$$G_+ = e^{-ABr}\left[b_0 r^{\gamma} + b_1 r^{\gamma+1} + \cdots + b_s r^{\gamma+s} + \cdots\right] \tag{23}$$

dans lesquelles l'exposant γ doit en principe être supposé positif pour que F_+ et G_+ restent finies pour $r=0$. En substituant (23) dans (18) et en annulant les coefficients de $r^{\gamma-1}$ on obtient :

$$\alpha\, a_0 + \gamma\, b_0 - l b_0 = 0\,; \qquad -\alpha\, b_0 + \gamma\, a_0 + (l+2)\, a_0 = 0. \tag{24}$$

Les deux équations (24) ne sont compatibles que si leur déterminant est nul, ce qui nous donne la condition :

$$\alpha^2 + (\gamma - l)(\gamma + l + 2) = 0 \tag{25}$$

En résolvant (25) par rapport à γ, on trouve :

$$\gamma = -1 + \sqrt{1 - [\alpha^2 - l\,(l+2)]} = -1 + \sqrt{(l+1)^2 - \alpha^2}. \tag{26}$$

Nous conservons seulement le signe + devant le radical, car le signe — donnerait une valeur négative inacceptable pour γ. Si nous faisons $l=0$ dans l'expression (26), nous trouvons pour γ une valeur très petite négative. Il pourrait sembler que cette valeur est à rejeter puisqu'elle correspond à des fonctions F_+ et G_+ (et par suite à des fonctions d'onde Ψ_k) qui deviennent infinis pour $r=0$. Or les résultats aux-

quels nous allons parvenir plus loin montrent clairement que l'on doit conserver cette solution avec $l = 0$.

On peut expliquer cette anomalie en remarquant que les fonctions d'onde Ψ_k deviennent pour $r = 0$ infinies d'un ordre si petit (à cause de la petitesse de α^2) que les intégrales $\int\int\int_{-\infty}^{+\infty} |\Psi_k|^2 d\tau$ restent convergentes. La condition que les fonctions d'onde soient partout finies apparaît donc ici comme trop rigoureuse: la condition essentielle, c'est qu'elles soient de carré sommable. Nous avons d'ailleurs vu que la même difficulté existait dans la théorie relativiste à une fonction d'onde (Cf. Chap. VIII, paragr. 2). Nous adopterons donc comme valeurs possibles de γ les valeurs (26) avec $l = 0, 1, 2, \ldots$

Ceci posé, revenons à la substitution de (23) dans (18) et annulons le coefficient de $r^{\gamma+s}$ nous trouvons :

$$\begin{aligned} A(A\,a_s - B\,b_s) + \alpha a_{s+1} + (\gamma + s + 1 - l)\,b_{s+1} &= 0 \\ -B(A\,a_s - B\,b_s) - \alpha\,b_{s+1} + (\gamma + s + l + 3)\,a_{s+1} &= 0. \end{aligned} \tag{27}$$

Multiplions la première équation (27) par B, la seconde par A et ajoutons. Il vient :

$$\begin{aligned} a_{s+1}\,[B\,\alpha + A(\gamma + s + l + 3)] &+ \\ + b_{s+1}\,[B(\gamma + s - l + 1) - A\,\alpha] &= 0 \end{aligned} \tag{28}$$

ce qui permet de poser :

$$\begin{aligned} a_{s+1} &= c_{s+1}\,[B(\gamma + s - l + 1) - A\,\alpha] \\ b_{s+1} &= -\,c_{s+1}\,[B\,\alpha + A(\gamma + s + l + 3)] \end{aligned} \tag{29}$$

où c_{s+1} est une certaine constante. En substituant les relations (29) et celles qui s'en déduisent en changeant $s+1$ en s, dans les équations (27), on trouve pour les c_s la relation de récurrence :

$$\begin{aligned} A\,c_s\,[-\alpha(A^2 - B^2) + 2\,AB(\gamma + s + 1)] &= \\ = c_{s+1}\,[A\,\alpha^2 + A(\gamma + s + l + 3)(\gamma + s - l + 1)]. \end{aligned} \tag{30}$$

Pour que les fonctions F_+ et G_+ soient sûrement nulles à l'infini, il suffit d'après (23) que les séries $a_0 r + \ldots$ et $b_0 r + \ldots$

soient limitées. Pour cela il faut que pour une certaine valeur de s, soit $s=p$, c_{s+1} soit nul, c_s ne l'étant pas ; il faut donc que pour $s=p$, le coefficient de c_s dans (30) soit nul ce qui nous donne :

$$\gamma + p + 1 = \alpha \frac{A^2 - B^2}{2\,AB} \tag{31}$$

ou d'après (26) :

$$\sqrt{(l+1)^2 - \alpha^2} + p = \alpha \frac{A^2 - B^2}{2\,AB}. \tag{32}$$

Mais, si nous nous reportons aux définitions (17), nous voyons que :

$$\frac{A^2 - B^2}{2\,AB} = \frac{W}{\sqrt{m^2_0 c^4 - W^2}} = \frac{E + m_0 c^2}{\sqrt{-2\,m_0 c^2\,E - E^2}} = \frac{1 + \frac{E}{m_0 c^2}}{\sqrt{-\frac{2\,E}{m_0 c^2}\left(1 + \frac{E}{2\,m_0 c^2}\right)}} \tag{33}$$

où $E = W - m_0 c^2$ est ici négatif. L'équation (32) peut donc s'écrire en élevant au carré :

$$\left[\sqrt{(l+1)^2 - \alpha^2} + p\right]^2 = \alpha^2 \frac{\left(1 + \frac{E}{m_0 c^2}\right)^2}{-\frac{2\,E}{m_0 c^2}\left(1 + \frac{E}{2\,m_0 c^2}\right)} \tag{34}$$

ou aussi bien :

$$\frac{-\frac{2\,E}{m_0 c^2}\left(1 + \frac{E}{2\,m_0 c^2}\right)}{\left(1 + \frac{E}{m_0 c^2}\right)^2} = \frac{\alpha^2}{\left[p + \sqrt{(l+1)^2 - \alpha^2}\right]^2} \tag{35}$$

Si nous ajoutons 1 aux deux membres de (35), nous trouvons :

$$\frac{1}{\left(1 + \frac{E}{m_0 c^2}\right)^2} = 1 + \frac{\alpha^2}{\left(p + \sqrt{(l+1)^2 - \alpha^2}\right)^2} \tag{36}$$

d'où la formule finale :

$$1+\frac{E}{m_0 c^2}=\left[1+\frac{\alpha^2}{(p+\sqrt{(l+1)^2-\alpha^2})^2}\right]^{-\frac{1}{2}} \tag{37}$$

Comme pour les solutions du type (I) dont nous nous occupons en ce moment, on a $j=l+\frac{1}{2}$, la formule (37) peut aussi s'écrire :

$$1+\frac{E}{m_0 c^2}=\left[1+\frac{\alpha^2}{\left[p+\sqrt{\left(j+\frac{1}{2}\right)^2-\alpha^2}\right]^2}\right]^{-\frac{1}{2}} \tag{38}$$

La formule (38) est analogue à la formule de Sommerfeld (¹), le nombre entier $j+\frac{1}{2}=l+1$ jouant le rôle de nombre quantique azimutal et p celui de nombre quantique radial.

Comme on le voit, les niveaux d'énergie du type (I) sont complètement définis par les 4 nombres quantiques : l, $j(=l+\frac{1}{2})$, m et p. Au lieu du nombre entier p qui peut prendre les valeurs 0, 1, ..., on peut tout aussi bien envisager le nombre $n=p+l+1$ qui peut prendre les valeurs 1, 2... Le nombre n est le « nombre quantique total » et l'ensemble des nombres n, l, j, m fixe la solution envisagée.

Si l'on développe le second membre de (38) jusqu'au second ordre en α^2, on obtient la formule approchée :

$$E=-\frac{Rh}{n^2}\left[1+\frac{\alpha^2}{n^2}\left(\frac{n}{j+\frac{1}{2}}-\frac{3}{4}\right)\right] \tag{39}$$

tout à fait comparable à la formule approchée de Sommerfeld (²) mais avec $j+\frac{1}{2}$ à la place du nombre k.

Rappelons que la Mécanique ondulatoire relativiste à une fonction Ψ nous avait conduit à une expression en désaccord

---

(¹) Cf. formule (38) du chapitre I.
(²) Cf. formule (41) du chapitre I.

avec les faits expérimentaux, expression qui avait même forme que (38), mais avec l au lieu de j [1].

b) *Solutions du type (II).*

Nous pourrions recommencer les calculs qui viennent d'être faits en partant cette fois des équations auxquelles satisfont les fonctions F_- et G_-. Mais cela est inutile car nous avons remarqué qu'on obtenait ces équations en changeant dans les équations (16) l en $-(l+1)$. Il est donc évident qu'en adoptant pour F_- et G_- des développements de la forme (23), nous obtiendrons pour déterminer γ l'équation qui se déduit de (26) par ledit changement de l en $-(l+1)$, c'est-à-dire

$$\gamma = -1 + \sqrt{l^2 - \alpha^2} \tag{40}$$

Ici l'on doit évidemment exclure le cas $l = 0$ qui donnerait une valeur imaginaire : nous savions d'ailleurs déjà qu'il n'y avait pas de solution du type (II) pour $l = 0$. Si $l = 1$, la formule (40) nous donne une valeur très petite négative de γ : ici encore nous admettrons cette valeur de l en remarquant que si les fonctions d'onde correspondantes sont infinies d'ordre peu élevé à l'origine des coordonnées, elles sont néanmoins de carré sommable. Bref nous admettrons pour l les valeurs possibles : 1, 2...

En poursuivant les calculs, nous parviendrons évidemment à une formule qui se déduit de (37) par la substitution de $-(l+1)$ à l, savoir :

$$1 + \frac{E}{m_0 c^2} = \left[1 + \frac{\alpha^2}{(p + \sqrt{l^2 - \alpha^2})^2}\right]^{-\frac{1}{2}} \tag{41}$$

Mais pour les solutions du type (II), on a : $j = l - \frac{1}{2}$. Nous retombons donc sur la formule (38).

Au lieu de caractériser une solution du type (II) par les 4 nombres quantiques l, $j\left(= l - \frac{1}{2}\right)$, m, p, nous pouvons

(1) Cf. formules (34) et (36) du chapitre VIII.

tout aussi bien la caractériser par les 4 nombres n, l, j, m avec $n = p + l$ et nous retrouvons facilement la formule approchée (39).

En résumé, dans tous les cas, l'énergie du niveau caractérisé par les nombres quantiques (n, l, j, m) est donnée (au second ordre en α^2) par la formule (39) (1).

Nous retrouvons donc pour l'hydrogène la formule de structure fine de Sommerfeld, mais avec cette différence essentielle que le nombre azimutal k de l'ancienne théorie des quanta y est remplacé par le nombre entier $j + \frac{1}{2}$. C'est donc entre les niveaux dont les nombres j diffèrent d'une unité que l'on doit trouver les doublets de Sommerfeld. Les niveaux qui ne diffèrent que par le nombre quantique azimutal (k ou l) sont confondus. Ceci est parfaitement d'accord avec la structure fine réelle des raies de la série de Balmer, par exemple de H_α ; on le verra en se reportant au paragr. 5 du chapitre III et en particulier à la figure 4.

On peut d'ailleurs reprendre rigoureusement toute la théorie précédente pour un atome hydrogénoïde (atome de numéro atomique N ionisé N — 1 fois). On trouve alors aisément au lieu de (39) :

$$E = -\frac{RhN^2}{n^2}\left[1 + \frac{\alpha^2 N^2}{n^2}\left(\frac{n}{j+\frac{1}{2}} - \frac{3}{4}\right)\right] \qquad (42)$$

Cette formule rend bien compte de la structure fine du spectre de He^+.

4. *Application des résultats obtenus aux spectres Röntgen.*

— Les formules fournies par la théorie de Dirac vont nous permettre d'interpréter la structure des spectres de Rayons X et, en particulier, l'existence et la valeur des doublets réguliers sans tomber dans les difficultés que rencontrait l'an-

(1) Notons qu'il y a dégénérescence, le nombre m n'intervenant pas dans l'expression de l'énergie quantifiée.

cienne théorie de Sommerfeld. Nous avons exposé au chapitre III les principaux faits expérimentaux que la théorie a à interpréter ; le lecteur pourra s'y reporter.

Nous avons vu que, pour calculer les niveaux d'énergie dans les atomes complexes, on peut tenir compte, assez grossièrement, de l'interaction des électrons en introduisant un « nombre d'écran », c'est-à-dire en remplaçant dans les formules qui sont valables pour les atomes hydrogénoïdes le nombre atomique N par un nombre réduit $N - z$. Naturellement le nombre d'écran z varie d'un électron à l'autre, étant en gros d'autant plus grand que l'électron est plus périphérique. Guidés par l'ancienne image de l'atome de Bohr, nous pouvons admettre que le nombre d'écran pour un électron intra-atomique dépend seulement des nombres quantiques n et l relatifs à cet électron. En nous servant de la formule (42), nous écrirons donc que, dans un atome complexe, l'électron de nombres quantiques n, l, j, m a une énergie donnée par la formule approximative :

$$E(n, l, j, m) = \\ -\frac{Rh(N - z_{n,l})^2}{n^2}\left[1 + \frac{\alpha^2 (N - z_{n,l})^2}{n^2}\left(\frac{n}{j + \frac{1}{2}} - \frac{3}{4}\right)\right]. \qquad (43)$$

Cette formule diffère de l'ancienne formule de Sommerfeld [cf. Chap. III, form. (9)] par la substitution de $j + \frac{1}{2}$ à k. Mais cette simple substitution suffit pour que la formule (43) permette d'interpréter les doublets réguliers sans prêter aux objections soulevées par la formule ancienne de Sommerfeld. Il est en effet évident qu'avec la formule (43) les doublet réguliers vont être prévus non pas entre niveaux de nombres quantiques azimutaux (k ou l) différant d'une unité, mais entre niveaux de même nombre quantique azimutal et de nombre j différant d'une unité : c'est bien là ce qu'exige l'expérience, nous l'avons vu.

Considérons par exemple les niveaux L_{II} et L_{III} dont la différence de fréquence correspond à des doublets de Sommerfeld. Nous savons que l'on a :

$$\text{pour } L_{II} \quad n=2 \quad l=1 \quad j=\frac{1}{2}$$

$$\text{pour } L_{III} \quad n=2 \quad l=1 \quad j=\frac{3}{2}.$$

Nous trouvons donc d'après (43) pour la différence de fréquence des doublets correspondants :

$$\nu_{L_{II}} - \nu_{L_{III}} = -\alpha^2 \frac{R(N-z_{21})^4}{2^4}\left[\frac{2}{\frac{1}{2}+\frac{1}{2}} - \frac{2}{\frac{3}{2}+\frac{1}{2}}\right] =$$

$$= -\frac{R\alpha^2}{16}(N-z_{21})^4 \qquad (44)$$

et nous avons vu qu'en posant $z=3{,}5$, on a un bon accord avec l'expérience. Et ici le doublet est prévu à sa vraie place entre deux niveaux de même l et de j différant d'une unité.

Nous avons signalé qu'il existait aussi dans les spectres X des doublets irréguliers. Ces doublets ont pour origine la différence de fréquence qui existe entre deux niveaux voisins de même nombre quantique j dont les nombres l diffèrent d'une unité. Nous pouvons aussi en expliquer l'existence grâce à la formule (43). En effet si le nombre d'écran z ne dépendait pas de l, les niveaux de mêmes j et de l différents (tels que L_I et L_{II} par exemple) seraient confondus : c'est ce qui se produit pour l'hydrogène car en ce cas, bien entendu, z est toujours nul. Mais en raison de la variation de z avec l, on voit que ces niveaux dans les atomes complexes ne doivent pas être confondus et on peut même prévoir une loi pour leur différence de fréquence. On aura en effet si l'on néglige les termes d'ordre supérieur en α^2 :

$$\sqrt{\frac{E(n,l,j)}{h}} = \sqrt{\frac{Rh}{n^2}}\,(N-z_{n,l})\,;$$

$$\sqrt{\frac{E(n,l+1,j)}{h}} = \sqrt{\frac{Rh}{n^2}}\,(N-z_{n,l+1}) \qquad (45)$$

D'où :

$$\sqrt{\frac{\mathrm{E}(n, l, j)}{h}} - \sqrt{\frac{\mathrm{E}(n, l+1, j)}{h}} = \sqrt{\frac{\mathrm{R}h}{n^2}}\,(z_{n,\,l+1} - z_{n,\,l}) \tag{46}$$

Le premier membre de (46) est $\delta\sqrt{\nu}$ pour les deux niveaux, le second est indépendant de N. On voit donc que les termes spectraux des deux niveaux dans la série des éléments sont tels que $\delta\sqrt{\nu}$ soit indépendant de N : c'est bien la loi des doublets irréguliers qui a été signalée à la fin du paragr. 2 du chapitre III. Ces doublets irréguliers existent entre termes spectraux caractérisés par la même lettre avec des indices différant d'une unité, *le premier étant impair*, tandis que les doublets réguliers de Sommerfeld proviennent des termes spectraux caractérisés par la même lettre avec des indices différant d'une unité, *le premier étant pair*. Ainsi la différence $\nu_{\mathrm{L_I}} - \nu_{\mathrm{L_{II}}}$ donne naissance à des doublets irréguliers et la différence $\nu_{\mathrm{L_{II}}} - \nu_{\mathrm{L_{III}}}$ à des doublets réguliers.

La théorie de Dirac a donc entièrement remis dans la bonne forme la théorie de structure fine de Sommerfeld montrant que le succès primitif n'était pas dû au hasard et qu'en dernière analyse, l'existence des doublets réguliers est bien reliée, par l'intermédiaire du spin de l'électron, à la Relativité.

5. ***Nombres d'électrons par niveaux: règle de Stoner.*** — La théorie de Dirac n'a pas seulement permis de retrouver en les améliorant les formules de structure fine, elle a aussi fourni une interprétation de la règle de Stoner relative à la répartition des électrons entre les niveaux de l'atome et une démonstration des règles de sélection relatives aux nombres quantiques l, m et j.

Occupons-nous d'abord ici de la première question et rappelons que la règle de Stoner est la suivante : « Il ne peut pas y avoir plus de $2j+1$ électrons sur le niveau d'énergie correspondant aux nombres quantiques n, l, j. » Nous allons chercher à justifier cet énoncé.

Nous avons vu plus haut (paragr. 2) que pour les solutions du type (I) comme pour les solutions du type (II), il existait $2j+1$ niveaux correspondant à l'ensemble des nombres quantiques (n, l, j). En d'autres termes, chaque niveau étant entièrement défini par les 4 nombres quantiques n, l, j, m, il existe $2j+1$ valeurs possibles de m et par suite $2j+1$ niveaux possibles pour des valeurs données des 3 nombres $n, l, j = l \pm \frac{1}{2}$. Ceci rappelé, pour aller plus loin, il nous faut faire appel à un principe nouveau qui joue un rôle fondamental dans la physique contemporaine : le principe d'exclusion de Pauli.

Nous énoncerons ici le principe d'exclusion de Pauli en disant : « Il ne peut y avoir plus d'un électron intra-atomique dont l'état stationnaire soit caractérisé par les quatre mêmes nombres entiers n, l, j, m. »

Or, nous l'avons vu, dans l'atome en l'absence de champ extérieur, l'énergie d'un état stationnaire ne dépend pas du nombre m de sorte que chaque niveau d'énergie est entièrement caractérisé par les 3 nombres quantiques n, l, j comme l'expérience du reste l'avait indiqué avant la théorie (cf. première partie). Si donc on admet le principe de Pauli, le nombre des électrons appartenant à un même niveau (n, l, j) est au plus égal au nombre des solutions stationnaires caractérisées par les 4 nombres quantiques n, l, j, m, dont les trois premiers ont les valeurs qui caractérisent le niveau considéré. Nous savons que ce nombre de solutions est $2j+1$. Nous avons donc justifié la règle de Stoner dont l'exactitude expérimentale ne paraît pas douteuse.

6. *Les règles de sélection en théorie de Dirac.* — Nous avons expliqué au paragraphe 4 du chapitre VIII comment le principe de correspondance conduit en Mécanique ondulatoire à la prévision des règles de sélection. Dans la théorie de Dirac, cette prévision se fait de la même façon, mais en tenant compte de ce fait que l'expression $\Psi\Psi^*$ de la Méca-

nique ondulatoire à une fonction d'onde est ici remplacée par

$$\sum_{1}^{4} {}_k \Psi_k^* \Psi_k \, .$$

On aura ainsi à considérer comme élément de la matrice X qui sert à évaluer la probabilité des transitions accompagnées de rayonnement, les quantités suivantes :

$$X_{nm} = \int_{-\infty}^{-\infty}\int\int x\,(\Psi_{1,n}^* \Psi_{1,m} + \Psi_{2,n}^* \Psi_{2,m}$$
$$+ \Psi_{3,n}^* \Psi_{3,m} + \Psi_{4,n}^* \Psi_{4,m})\, d\tau \, . \quad (47)$$

où les indices n et m caractérisent deux états stationnaires et représentent en réalité chacun un ensemble de 4 indices n, l, j, m. Si les éléments X_{nm} Y_{nm} Z_{nm} sont nuls, il n'y a pas de rayonnement correspondant au passage $n \longrightarrow m$. De là, on tire les règles de sélection.

Comme exemple, considérons la transition depuis l'état stationnaire du type (I) de nombres quantiques n, l, $j = l + \frac{1}{2}$, m jusqu'à l'état stationnaire du type (II) de nombres quantiques n, $l-1$, $j = l - \frac{3}{2}$, m ; transition pour laquelle on a $|\delta l| = 1$, $\delta m = 0$ et $|\delta j| = 2$. Le premier de ces états a pour fonctions d'onde :

$$\Psi_1 = i\, F_+ (r)\, Y_{l+1}^{m} \, ; \quad \Psi_2 = -\, i\, F_+ (r)\, Y_{l-1}^{m-1} \, ;$$
$$\Psi_3 = (l - m + 1)\, G_+ (r)\, Y_l^{m} ; \quad \Psi_4 = (l + m)\, G_+ (r)\, Y_l^{m-1}$$

et le second :

$$\Psi_1 = -\, i\, (l - m - 1)\, F_- (r)\, Y_{l-2}^{m}$$
$$\Psi_2 = i\, (l + m - 2)\, F_- (r)\, Y_{l-2}^{m-1} ;$$
$$\Psi_3 = G_- (r)\, Y_{l-1}^{m} \, ; \quad \Psi_4 = -\, G_- (r)\, Y_{l-1}^{m-1} .$$

Ecrivons alors l'expression de l'élément de la matrice Z

correspondant à cette transition, en posant pour simplifier :

$$A = -\int_0^{+\infty} F_+(r)\, F_-(r)\, r^2\, dr\ ;\quad B = \int_0^{+\infty} G_+(r)\, G_-(r)\, r^2\, dr. \tag{48}$$

Il vient :

$$\begin{aligned} Z_{nm} = {} & A\,(l - m - 1) \int\!\!\int \cos\theta\, \overset{m*}{\underset{l-2}{Y}}\, \overset{m}{\underset{l+1}{Y}}\, d\Omega + \\ & + A\,(l + m - 2) \int\!\!\int \cos\theta\, \overset{m-1\,*}{\underset{l-2}{Y}}\, \overset{m-1}{\underset{l+1}{Y}}\, d\Omega + \\ & + B\,(l - m + 1) \int\!\!\int \cos\theta\, \overset{m*}{\underset{l-1}{Y}}\, \overset{m}{\underset{l}{Y}}\, d\Omega - \\ & - B\,(l + m) \int\!\!\int \cos\theta\, \overset{m-1}{\underset{l-1}{Y}}\, \overset{m-1\,*}{\underset{l}{Y}}\, d\Omega \end{aligned} \tag{49}$$

avec $d\Omega = \sin\theta\, d\theta\, d\varphi$, les intégrales doubles étant étendues à la surface de la sphère de rayon $u\,n$. Or, en partant de la définition des fonctions Y, on peut démontrer les formules suivantes :

$$\int\!\!\int \cos\theta\, \overset{m*}{\underset{l-2}{Y}}\, \overset{m}{\underset{l+1}{Y}}\, d\Omega = 0\ ;$$

$$\int\!\!\int \cos\theta\, \overset{m*}{\underset{l-1}{Y}}\, \overset{m}{\underset{l}{Y}}\, d\Omega = \frac{4\pi}{(2l+1)(2l-1)}\,(l+m)!\,(l-m)!$$

$$\int\!\!\int \cos\theta\, \overset{m-1\,*}{\underset{l-2}{Y}}\, \overset{m-1}{\underset{l+1}{Y}}\, d\Omega = 0\ ;$$

$$\int\!\!\int \cos\theta\, \overset{m-1\,*}{\underset{l-1}{Y}}\, \overset{m-1}{\underset{l}{Y}}\, d\Omega = \frac{4\pi}{(2l+1)(2l-1)}\,(l+m-1)!\,(l-m+1)!$$

Portons dans (49) ; nous trouvons :

$$Z_{nm} = B\,\frac{4\pi}{(2l+1)(2l-1)} \Big[(l-m+1)\,(l+m)!\,(l-m)! - (l+m)\,(l-m+1)!\,(l+m-1)!\Big] = 0\,. \tag{51}$$

On trouverait de même $X_{nm} = Y_{nm} = 0$. La transition envisagée est donc interdite.

On peut recommencer le même genre de calculs en prenant toutes les combinaisons d'une solution du type (I) avec une solution du même type ou du type (II), puis toutes les combinaisons d'une solution du type (II) avec une solution du même type ou du type (I). Le résultat de ces calculs est le suivant :

« Seules correspondent à un rayonnement les transitions pour lesquelles on a :

$$\delta l = \pm 1 \qquad \delta m = \begin{cases} \pm 1 \\ 0 \end{cases} \qquad \delta j = \begin{cases} \pm 1 \\ 0 \end{cases} \text{. »}$$

Ces règles de sélection sont bien celles que l'on a trouvées empiriquement. Les règles relatives aux nombres quantiques *l* et *m* étaient prévues par la Mécanique ondulatoire à une seule fonction d'onde, mais la règle relative au nombre quantique *j* ne pouvait l'être que par une théorie qui introduit ce nombre quantique, comme c'est le cas pour la théorie de Dirac.

CHAPITRE XVIII

Déduction de la formule de Landé

---

1. Vue sommaire sur la méthode des perturbations

NOUS nous proposons de montrer dans ce chapitre que la théorie de Dirac permet de retrouver la formule de Landé pour l'effet Zeeman anomal des alcalins dans un champ magnétique faible. Mais, comme nous devons pour cela faire usage de la méthode des perturbations, nous allons d'abord dire quelques mots de cette méthode de calcul.

Supposons que nous ayons déterminé les états stationnaires d'un système, d'un atome d'hydrogène par exemple. Nous connaissons donc les valeurs propres W_n de l'énergie et les fonctions propres correspondantes Ψ_{kn}, k désignant toujours l'indice de Dirac, la variable de « spin ». L'équation symbolique

$$\left[\frac{W_n + eV}{c} + \sum_{1}^{3} \alpha_j \; P_j + \alpha_4 \, m_0 \, c\right] \Psi_n = 0 \qquad (1)$$

est donc satisfaite quand le système n'est pas perturbé.

Mais supposons maintenant qu'une légère action perturbatrice constante s'exerce sur le système quantifié et qu'elle peut être représentée par l'addition d'un terme $\Lambda\Psi$ à l'équation (1). (Λ est un opérateur qui peut contenir les α_i et par

suite opérer sur l'indice k.) Par suite de la présence du petit terme perturbateur, les valeurs et fonctions propres vont être légèrement modifiées et devenir $W_n + \varepsilon_n$ et $\Psi_{k,n} + \eta_{k,n}$.

Nous supposerons que les ε_n et les $\eta_{k,n}$ sont très petits comme le terme perturbateur $A\,\Psi$ et nous négligerons les termes tels que $\varepsilon_n\,\eta_{k,n}$ et $A\,\eta_{k,n}$.

Dans l'état perturbé, l'équation symbolique est :

$$\left[\frac{W_n + \varepsilon_n + eV}{c} + \sum_{1}^{3} \alpha_j\,P_j + \alpha_4\,m_0c + A\right](\Psi_n + \eta_n) = 0. \quad (2)$$

Retranchons (1) de (2); nous trouvons avec les approximations admises :

$$\left(\frac{\varepsilon_n}{c} + A\right)\Psi_n + \left[\frac{W_n + eV}{c} + \sum_{1}^{3} \alpha_j\,P_j + \alpha_4\,m_0c\right]\eta_n = 0. \quad (3)$$

Considérons maintenant les 4 fonctions $\eta_{k,n}$ il résulte de la formule (5) du chapitre XVI que nous pouvons développer ces fonctions suivant le système complet des fonctions Ψ_k par les formules :

$$\eta_{k,n} = \sum_m c_{n,m}\,\Psi_{k,m} \quad (k = 1, 2, 3, 4) \quad (4)$$

ou symboliquement :

$$\eta_n = \sum_m c_{n,m}\,\Psi_m \quad (5)$$

On a donc en tenant compte de l'équation obtenue en changeant n en m dans (1) :

$$\left(\frac{W_n + eV}{c} + \sum_{1}^{3} \alpha_j\,P_j + \alpha_4\,m_0c\right)\eta_n =$$

$$= \sum_m c_{n,m}\left[\frac{W_n + eV}{c} + \sum_{1}^{3} \alpha_j\,P_j + \alpha_4\,m_0c\right]\Psi_m$$

$$= \sum_m c_{n,m}\,\frac{W_n - W_m}{c}\,\Psi_m \quad (6)$$

et, par suite, d'après (3) :

$$\sum_m c_{n,m} \frac{W_n - W_m}{c} \Psi_m = \left(\frac{\varepsilon_n}{c} + \Lambda\right) \Psi_n \qquad (7)$$

ou en explicitant :

$$\sum_m c_{n,m} \frac{W_n - W_m}{c} \Psi_{k,m} = \left(\frac{\varepsilon_n}{c} + \Lambda\right) \Psi_{k,n}. \qquad (8)$$

Multiplions dans (8) par $\Psi^*_{k,m}$, sommons par rapport à k et intégrons dans l'espace. Il vient :

$$\sum_m c_{n,m} \frac{W_n - W_m}{c} \int\int_{-\infty}^{+\infty}\int \sum_k \Psi^*_{k,l} \Psi_{k,m} \, d\tau =$$

$$= \int\int_{-\infty}^{+\infty}\int \sum_k \Psi^*_{k,l} \left(\frac{\varepsilon_n}{c} + \Lambda\right) \Psi_{k,n} \, d\tau . \qquad (9)$$

d'où l'on tire, en tenant compte de l'orthogonalité et de la normalisation des fonctions propres :

$$c_{n,l} = \frac{c}{W_n - W_l} \int\int_{-\infty}^{+\infty}\int \sum_k \Psi^*_{k,l} \left(\frac{\varepsilon_n}{c} + \Lambda\right) \Psi_{k,n} \, d\tau . \qquad (10)$$

Cette formule nous donne les c_{nl} pour toutes les valeurs de l.

Mais la formule (10) nous donnerait pour $l = n$ un coefficient c_{nn} infini si l'intégrale qui figure dans (10) était différente de 0 pour $l = n$. Comme ceci n'est pas acceptable, nous voyons qu'il faut avoir :

$$\int\int_{-\infty}^{+\infty}\int \sum_k \Psi^*_{k,n} \left(\frac{\varepsilon_n}{c} + \Lambda\right) \Psi_{k,n} \, d\tau = 0. \qquad (11)$$

Cette condition (11) qui exprime un théorème bien connu de Fredholm dans la théorie des équations intégrales, permet d'obtenir la variation ε_n de l'énergie du n^e état stationnaire provoquée par la présence de la perturbation. De (11), on tire en effet :

$$\frac{\varepsilon_n}{c} = -\int\int_{-\infty}^{+\infty}\int \Sigma_k \Psi_{k,n}^* \Lambda \Psi_{k,n} \, d\tau \qquad (12)$$

parce que les $\Psi_{k,n}$ sont normées.

Il faut remarquer que la formule (10) n'est acceptable que si elle conduit à de petites valeurs des c_{nl} ; sans quoi, l'hypothèse de la petitesse des η_n, sur laquelle nous nous appuyons, ne serait pas valable. En examinant ce point, l'on voit que la formule (12) n'est exacte que si l'on peut, comme nous l'avons implicitement supposé, considérer la perturbation de chaque valeur propre comme indépendante de celles subies par les autres valeurs propres. Pour cela, la perturbation extérieure doit être assez faible pour que le déplacement des valeurs propres dû à la perturbation soit petit devant la différence de ces valeurs propres : c'est ce qui a lieu pour l'effet Zeeman quand le champ magnétique est faible. Si cette condition n'est pas remplie (effet Zeeman avec champ magnétique fort, par exemple), il y a lieu de diriger un peu autrement le calcul des perturbations : nous n'y insisterons pas ici.

2. *Application de la formule* (12) *dans le cas de l'effet Zeeman.* — Nous avons vu au Chapitre IV comment la formule de Landé permet de représenter la modification d'un niveau d'énergie dans un atome à spectre de doublets (alcalins) par l'action d'un faible champ magnétique. Si $W_0(n, l, j)$ est l'énergie du niveau en l'absence de champ magnétique et si $W_H(n, l, j)$ est cette énergie en présence du champ H, on a :

$$W_H(n,l,j) = W_0(n, l, j) + m'g\frac{eh\,H}{4\pi m_0 c} \tag{13}$$

avec :

$$g = \frac{j+\frac{1}{2}}{l+\frac{1}{2}} = \frac{2j+1}{2l+1}. \tag{14}$$

Le nombre m' est un nombre demi-entier (c'est-à-dire un entier plus $\frac{1}{2}$) positif ou négatif pouvant prendre toutes les valeurs demi-entières de $-j$ à $+j$. Nous avons modifié les

notations du chapitre IV en écrivant m' au lieu de m afin de distinguer le nombre demi-entier m' de la formule (13) du quatrième nombre quantique m caractérisant un état stationnaire de l'atome, nombre qui, lui, est entier. Nous verrons que les nombres m' et m sont unis par la relation $m' = -\left(m - \frac{1}{2}\right)$. Précisons ce point : dans la formule (13), nous définissons le niveau d'énergie par ses trois nombres quantiques n, l, j, ce qui suffit puisque la valeur de l'énergie ne dépend pas du quatrième nombre quantique m ; mais le déplacement du niveau par effet Zeeman dépend de m [1] et, en explicitant, on peut écrire (13) sous la forme :

$$W_H(n,l,j,m) = W_0(n,l,j) - \left(m - \frac{1}{2}\right)\frac{2j+1}{2l+1}\frac{eh\mathrm{H}}{4\pi m_0 c} \quad (15)$$

C'est cette formule que la théorie de Dirac va nous permettre de retrouver.

Nous considérons un atome où il est permis de n'envisager qu'un seul électron : c'est le cas rigoureusement pour un atome d'hydrogène et pour un atome de rang N ionisé (N — 1) fois ; c'est approximativement le cas pour un atome alcalin où l'électron de valence est unique et où l'on peut tenir compte grossièrement de la présence des autres électrons périphériques par un simple effet d'écran. Nous supposerons que cet atome est plongé dans un champ magnétique uniforme H dirigé suivant l'axe des z. Nous pouvons alors prendre comme potentiel vecteur :

$$A_x = -\frac{1}{2} y \mathrm{H} \; ; \; A_y = \frac{1}{2} x \mathrm{H} \; ; \; A_z = 0 \quad (16)$$

Le terme perturbateur de l'équation (2) est donc :

$$\Lambda \Psi = \frac{e}{c}(\alpha_1 A_x + \alpha_2 A_y + \alpha_3 A_z)\Psi = \frac{e\mathrm{H}}{2c}(x\alpha_2 - y\alpha_1)\Psi \quad (17)$$

---

[1] On peut dire que la présence du champ magnétique fait disparaître la dégénérescence qui existait en son absence.

En explicitant grâce aux valeurs bien connues de α_1 et α_2, on trouve :

$$\Lambda \Psi_1 = i\,(x + i\,y)\,\frac{e\,H}{2\,c}\,\Psi_4\;;\; \Lambda \Psi_2 = -\,i\,(x - i\,y)\,\frac{e\,H}{2\,c}\,\Psi_3$$
$$\Lambda \Psi_3 = i\,(x + i\,y)\,\frac{e\,H}{2\,c}\,\Psi_2\;;\; \Lambda \Psi_4 = -\,i\,(x - i\,y)\,\frac{e\,H}{2\,c}\,\Psi_1 \qquad (18)$$

L'application de la formule (12) nous donne alors pour le déplacement du niveau par effet Zeeman dans un champ magnétique faible :

$$\frac{\varepsilon}{c} = \frac{W_H\,(n, l, j, m) - W_0\,(n, l, j)}{c}$$
$$= \frac{e\,H}{2c}\int\int\int_{-\infty}^{+\infty} [-\,\Psi_1{}^* i\,(x + i\,y)\,\Psi_4 + \Psi_2{}^* i\,(x - i\,y)\,\Psi_3$$
$$-\,\Psi_3{}^* i\,(x + i\,y)\,\Psi_2 + \Psi_4{}^* i\,(x - i\,y)\,\Psi_1]\,d\tau \qquad (19)$$

Tout le problème est maintenant de calculer l'intégrale qui figure en (19).

3. ***Développement du calcul.*** — Pour faire le calcul, nous devons distinguer deux cas suivant que l'état stationnaire d'énergie $W_0(n, l, j)$ est du type (I) ou du type (II).

a) *Solution du type (I):*

On a alors :

$$\Psi_1 = i\,F_+\,(r)\,Y_{l+1}^{m}\;;\; \Psi_2 = -\,i\,F_+\,(r)\,Y_{l+1}^{m-1},$$
$$\Psi_3 = (l - m + 1)\,G_+\,(r)\,Y_l^{m}\;;\; \Psi_4 = (l + m)\,G_+\,(r)\,Y_l^{m-1}. \qquad (20)$$

où F_+ et G_+ sont deux fonctions réelles de r obéissant aux équations simultanées :

$$\frac{2\,\pi}{h}\left[\frac{W_0 + e\,V}{c} + m_0\,c\right]F_+ + \frac{d\,G_+}{dr} - \frac{l}{r}\,G_+ = 0$$
$$-\frac{2\,\pi}{h}\left[\frac{W_0 + e\,V}{c} - m_0\,c\right]G_+ + \frac{d\,F_+}{dr} + \frac{l + 2}{r}\,F_+ = 0 \qquad (21)$$

Comme dans l'atome, l'approximation Newtonienne est toujours approximativement valable, les fonctions Ψ_1 et Ψ_2 sont très petites par rapport à Ψ_3 et Ψ_4 ; le rapport $\frac{F_\pm}{G_+}$ est donc très petit. Un calcul plus complet prouve que ce rapport est de l'ordre de la constante de structure fine α. Nous nous appuyerons sur ce fait pour négliger le carré $\left(\frac{F_+}{G_+}\right)^2$, ce qui apportera de notables simplifications.

Des équations (21), on tire, en multipliant la première par G_+, la seconde par F_+ et en ajoutant :

$$\frac{4\pi m_0 c}{h} F_+ G_+ + \frac{1}{2}\frac{d}{dr}\left[F_+{}^2 + G_+{}^2\right] + \frac{l+2}{r} F_+{}^2 - \frac{l}{r} G_+{}^2 = 0 \qquad (22)$$

Négligeons les termes en F_+, multiplions par $r^3\,dr$ et intégrons de 0 à $+\infty$; il vient :

$$\int_0^{+\infty} F_+ G_+ r^3\,dr = -\frac{h}{4\pi m_0 c}\int_0^{+\infty}\left[\frac{r^3}{2}\frac{d\,G^2{}_+}{dr} - l\,G_+{}^2 r^2\right] dr \qquad (23)$$

En intégrant par parties, on obtient, G_+ étant nul à l'infini :

$$\int_0^{+\infty} F_+ G_+ r^3\,dr = \frac{h}{4\pi m_0 c}\left(l + \frac{3}{2}\right)\int_0^{+\infty} G_+{}^2 r^2\,dr \qquad (24)$$

Munis de ces formules, nous allons pouvoir aborder le calcul de l'intégrale qui figure au second membre de (19). En prenant des coordonnées polaires, nous avons :

$$x + iy = r\sin\theta\, e^{i\varphi}\ ;\ x - iy = r\sin\theta\, e^{-i\varphi}\ ;$$
$$d\tau = r^2 \sin\theta\, dr\, d\theta\, d\varphi. \qquad (25)$$

Donc :

$$\int\int\int_{-\infty}^{+\infty}\Big[-\Psi_1{}^* i\,(x+iy)\,\Psi_4 + \Psi_2{}^* i\,(x-iy)\,\Psi_3$$
$$- \Psi_3{}^* i\,(x+iy)\,\Psi_2 + \Psi_4{}^* i\,(x-iy)\,\Psi_1\Big]\, d\tau =$$

$$= -\int_0^{2\pi}\int_0^{\pi}\Big[\overset{m*}{\underset{l+1}{Y}}(l+m)\overset{m-1}{\underset{l}{Y}}e^{i\varphi}+\overset{m-1*}{\underset{l+1}{Y}}(l-m+1)\,\overset{m}{Y_l}\,e^{-i\varphi}$$
$$+(l-m+1)\,\overset{m*}{Y_l}\;\overset{m-1}{\underset{+1}{Y}}e^{i\varphi}+(l+m)\overset{m-1*}{\underset{l}{Y}}\;\overset{m}{\underset{l+1}{Y}}e^{-i\varphi}\Big]$$
$$\sin^2\theta\, d\theta\, d\varphi\int_0^{+\infty}F_+\,G_+\,r^3\,dr. \tag{26}$$

Or l'étude des fonctions Y fournit les intégrales :

$$\int_0^{2\pi}\int_0^{\pi}\overset{\overset{*}{m}}{\underset{l+1}{Y}}\;\overset{m-1}{\underset{l}{Y}}e^{i\varphi}\sin^2\theta\,d\theta\,d\varphi=$$
$$=\frac{4\pi}{(2l+1)(2l+3)}(l+m+1)!\,(l-m+1)!$$

$$\int_0^{2\pi}\int_0^{\pi}\overset{\overset{*}{m-1}}{\underset{l+1}{Y}}\;\overset{m}{\underset{l}{Y}}\;e^{-i\varphi}\sin\theta\,d\theta\,d\varphi=$$
$$-\frac{4\pi}{(2l+1)\;(2l+3)}(l+m)!\,(l-m+2)!$$

$$\int_0^{2\pi}\int_0^{\pi}\overset{\overset{*}{m}}{\underset{l}{Y}}\;\overset{m-1}{\underset{l+1}{Y}}e^{i\varphi}\sin\theta\,d\theta\,d\varphi= \tag{27}$$
$$-\frac{4\pi}{(2l+1)\;(2l+3)}(l+m)!\,(l-m+2)!$$

$$\int_0^{2\pi}\int_0^{\pi}\overset{\overset{*}{m-}}{Y}\;\overset{m}{\underset{l+1}{Y}}\;e^{-i\varphi}\sin\theta\,d\theta\,d\varphi=$$
$$=\frac{4\pi}{(2l+1)\;(2l+3)}(l+m+1)!\,(l-m+1)!$$

On trouve donc pour l'intégrale (26) la valeur :

$$-\frac{4\pi}{(2l+1)\,(2l+3)}(l+m)!\,(l-m+1)!\,2.$$
$$\Big[(l+m)\,(l+m+1)-(l-m+1)\,(l-m+2)\Big]\int_0^{+\infty}F_+\,G_+\,r^3\,dr$$
$$=-\frac{4\pi}{2l+1}(l+m)!\,(l-m+1)!\,\Big[(l+m)\,(l+m+1)- \tag{28}$$
$$-(l-m+1)\,(l-m+2)\Big]\frac{h}{4\pi\, m_0 c}\int_0^{+\infty}G_+^{\,2}\,r^2\,dr$$

la dernière forme étant obtenue en introduisant l'expression (24).

Comme on vérifie aisément l'identité suivante :

$$(l+m)(l+m+1)-(l-m+1)(l-m+2) = (2l+2)(2m-1) \qquad (29)$$

on peut encore écrire l'expression (28) sous la forme :

$$-4\pi\cdot\frac{2l+2}{2l+1}(2m-1)(l+m)!\,(l-m+1)!\;\frac{h}{4\pi m_0 c}\int_0^{+\infty} G_+^2 r^2\,dr. \qquad (30)$$

Pour aller plus loin, il faut invoquer la condition de normalisation qui donne ici, en négligeant les termes en F_+^2 devant ceux en G_+^2 :

$$\int_0^{2\pi}\int_0^{\pi}\left[(l-m+1)^2\,\overset{*m}{Y_l}\,\overset{m}{Y_l}+(l+m)^2\,\overset{m-1}{Y_l^*}\,\overset{m-1}{Y_l}\right]\sin\theta\,d\theta\,d\varphi\int_0^{+\infty} G_+^2 r^2\,dr=1. \qquad (31)$$

Or on a :

$$\left.\begin{aligned}\int_0^{2\pi}\int_0^{\pi}\left|\overset{m}{Y_l}\right|^2\sin\theta\,d\theta\,d\varphi &= \frac{4\pi}{2l+1}(l+m)!\,(l-m)!\\ \int_0^{2\pi}\int_0^{\pi}\left|\overset{m-1}{Y_l}\right|^2\sin\theta\,d\theta\,d\varphi &= \frac{4\pi}{2l+1}(l+m-1)!\,(l-m+1)!\end{aligned}\right\} \qquad (32)$$

de sorte que la condition (31) nous donne :

$$\int_0^{+\infty} G_+^2 r^2\,dr = \frac{1}{4\pi(l+m)!\,(l-m+1)!}. \qquad (33)$$

En portant cette valeur dans l'expression (30), nous voyons finalement que l'intégrale à calculer de la formule (19) a simplement pour valeur :

$$-\frac{2l+2}{2l+1}\left(2m-1\right)\frac{h}{4\pi m_0 c}. \qquad (34)$$

La formule (19) nous fournit donc pour le déplacement du niveau par effet Zeeman, la formule :

$$W_H(n, l, j, m) = W_0(n, l, j) -$$
$$- \frac{eH}{2} \frac{2l+2}{2l+1} \left(2m - 1\right) \frac{h}{4\pi m_0 c}. \qquad (35)$$

Posons maintenant :

$$m' = -\left(m - \frac{1}{2}\right) \qquad g = \frac{l+1}{l+\frac{1}{2}} = \frac{j+\frac{1}{2}}{l+\frac{1}{2}} \qquad (36)$$

en nous souvenant qu'ici $j = l + \frac{1}{2}$. Nous pouvons écrire :

$$W_H(n, l, j, m) = W_0(n, l, j) + m'g \frac{ehH}{4\pi m_0 c} \qquad (37)$$

Si maintenant nous supposons seulement donnés les trois nombres quantiques n, l, j, nous aurons :

$$W_H(n, l, j) = W_0(n, l, j) + m'g \frac{ehH}{4\pi m_0 c} \qquad (38)$$

m' étant ici un nombre demi-entier qui peut varier de $-\left(m_{\min} - \frac{1}{2}\right)$ à $-\left(m_{\max} - \frac{1}{2}\right)$. Or, en discutant le nombre des niveaux correspondant au triplet de nombres quantiques (n, l, j), nous avons vu que pour une solution du type (I) on peut faire varier m de $-l$ à $l+1$. Le nombre m', pour (n, l, j) donné, peut donc varier de $-\left(-l - \frac{1}{2}\right)$ $= l + \frac{1}{2} = j$ à $-\left(l + 1 - \frac{1}{2}\right) = -\left(l + \frac{1}{2}\right) = -j$.

Bref, nous avons pour les niveaux correspondant à des solutions du type (I), la formule (38) où m' est un nombre demi-entier pouvant varier de $-j$ à $+j$. C'est bien la formule de Landé.

b) *Solution du type (II):*

Pour faire le calcul dans le cas d'une solution du type (II), on part des fonctions d'ondes :

$$\Psi_1 = -\,i\,(l - m)\,F_-\,Y_{l-1}^{m}\;;\quad \Psi_2 = i\,(l + m - 1)\,F_-\,Y_{l-1}^{m-1}\;;$$

$$\Psi_3 = G_-\,Y_{l}^{m}\;;\quad \Psi_4 = -\,G_-\,Y_{l}^{m-1} \tag{39}$$

les fonctions F_- et G_- étant des fonctions réelles qui satisfont aux équations :

$$\frac{2\pi}{h}\left[\frac{W + eV}{c} + m_0 c\right] F_- + \frac{dG_-}{dr} + \frac{l+1}{r}\,G_- = 0. \tag{40}$$

$$-\frac{2\pi}{h}\left[\frac{W + e\,V}{c} - m_0\,c\right] G_- + \frac{d\,F_-}{dr} - \frac{l-1}{2}\,F_- = 0$$

Par des calculs analogues à ceux développés plus haut, on parvient à la formule :

$$W_{II}\,(n, l, j, m) = W_0\,(n, l, j) - \frac{2\,l}{2\,l+1}\,(2\,m - 1)\,\frac{e\,H}{2}\cdot\frac{h}{4\,\pi\,m_0\,c} \tag{41}$$

Comme pour les solutions (II) on a $j = l - \frac{1}{2}$ nous poserons ici :

$$m' = -\left(m - \frac{1}{2}\right);\quad g = \frac{l}{l+\frac{1}{2}} = \frac{j+\frac{1}{2}}{l+\frac{1}{2}} \tag{42}$$

et nous pourrons écrire (41) sous la forme :

$$W_{II}\,(n, l, j, m) = W_0\,(n, l, j) + m'\,g\,\frac{e\,h\,H}{4\,\pi\,m_0\,c} \tag{43}$$

Supposons maintenant que, seuls, nous sont donnés les trois nombres quantiques n, l, j ; nous écrirons alors

$$W_{II}\,(n, l, j) = W_0\,(n, l, j) + m'\,g\,\frac{e\,h\,H}{4\,\pi\,m_0\,c} \tag{44}$$

m' étant un nombre demi-entier qui varie de $-\left(m_{\min} - \frac{1}{2}\right)$

à $-\left(m_{\text{max}} - \frac{1}{2}\right)$. Or ici (nous l'avons vu), on peut faire varier m depuis $-(l-1)$ jusqu'à $+l$. Le nombre m' peut donc varier de $-\left[-(l-1)-\frac{1}{2}\right]=l-\frac{1}{2}$ jusqu'à $-\left(l-\frac{1}{2}\right)$ c'est-à-dire de $+j$ à $-j$.

Bref nous avons pour les niveaux correspondants au type (II) la formule (44) où m' est un nombre demi-entier pouvant varier de $-j$ à $+j$.

4. ***Résumé des résultats; remarques.*** — Nous avons ainsi démontré que pour tous les niveaux, aussi bien du type (II) que du type (I), le déplacement du niveau (n, l, j) par l'action d'un champ magnétique uniforme suffisamment faible est donné par :

$$W_H(n, l, j) - W_0(n, l, j) = m' g \frac{e h H}{4 \pi m_0 c} \tag{45}$$

avec :

$$g = \frac{j + \frac{1}{2}}{l + \frac{1}{2}} \tag{46}$$

m' étant un nombre demi-entier pouvant prendre toutes les valeurs demi-entières entre $-j$ et $+j$. C'est bien là l'effet Zeeman anomal tel qu'il est exactement représenté par la formule de Landé. Le nombre m' ayant $2j+1$ valeurs possibles, tout niveau (n, l, j) est décomposé par l'effet Zeeman en $(2j+1)$ niveaux distincts, ce qui peut s'exprimer en disant que le champ magnétique extérieur a fait disparaître la dégénérescence d'ordre $2j+1$ qui existait en son absence.

Le raisonnement que nous avons fait est rigoureusement valable pour les atomes hydrogénoïdes, mais, comme nous l'avons indiqué, il peut s'appliquer au moins approximativement aux atomes alcalins.

Nous avons dit que la formule de Landé est valable pour les champs magnétiques faibles. Il faut entendre par là les

champs magnétiques assez faibles pour que le déplacement par effet Zeeman des niveaux d'énergie soit petit par rapport à la distance normale de ces niveaux. Si cette condition n'est pas remplie, la formule (12) n'est plus valable et il faut reprendre tout le calcul. On a pu démontrer (1) que l'on arrive alors à une formule identique à celle de Voigt (cf. Ch. IV, fin du parag. 3) et, pour les champs très forts, on trouve l'effet Paschen-Back, c'est-à-dire la décomposition Zeeman normale.

La théorie de Dirac donne donc une interprétation tout à fait satisfaisante de l'effet Zeeman anomal pour les alcalins. Pour les atomes où il n'est pas possible de considérer un seul électron optique, la théorie de Dirac ne peut prévoir rigoureusement l'effet Zeeman car cette théorie ne sait pas encore traiter le cas d'un système d'électrons en interaction.

(1) Voir C. G. Darwin (*Proc. Roy. Soc.*, vol 118 [1928], p. 654).

CHAPITRE XIX

Moment propre et moment orbital
Polarisation des ondes électroniques

1. *Impossibilité de séparer le moment magnétique propre du moment magnétique orbital (Bohr)*

M. Bohr a montré par des raisonnements assez délicats qu'il ne doit pas être possible de mesurer le moment magnétique propre d'un électron parce que les effets dus à l'existence de ce moment propre ne peuvent être distingués de ceux qui seraient dus au mouvement d'ensemble de l'électron.

Résumons le raisonnement de M. Bohr. Pour mettre en évidence le magnétisme propre de l'électron, on peut procéder de deux façons :

1° Chercher à mesurer l'action sur un magnétomètre du petit aimant équivalent à l'électron ;

2° Faire traverser par l'électron un champ magnétique non uniforme et chercher à observer l'action de ce champ sur le petit aimant magnétique.

Examinons la première méthode. La direction du mouvement de l'électron étant prise pour axe des x, nous plaçons le magnétomètre sur l'axe des y au point d'ordonnée y. Pour pouvoir écrire avec exactitude l'expression des actions sur le magnétomètre, nous devons supposer que l'électron

est suffisamment bien localisé, c'est-à-dire que cet électron doit être associé à un train d'ondes Ψ dont les dimensions soient petites par rapport à la distance y du magnétomètre à l'axe du mouvement. Par exemple, si Δx est la longueur de ce train d'ondes le long de l'axe des x (incertitude sur l'abscisse de l'électron), on doit avoir :

$$\Delta x \ll y \tag{1}$$

et de même :

$$\Delta y \ll y \tag{1'}$$

Ceci posé, en passant près de l'origine des coordonnées, l'électron va exercer deux actions sur le magnétomètre. L'une de ces actions est due au champ magnétique (champ orbital H_0) créé par le mouvement de translation de l'électron ; sa valeur est :

$$H_0 = \frac{e}{c}\frac{v_x}{y^2} \tag{2}$$

La seconde action est due au moment magnétique propre de l'électron qui crée au lieu occupé par le magnétomètre le champ magnétique propre :

$$H_P = \frac{e\,h}{4\pi m_0 c}\cdot\frac{1}{y^3} \tag{3}$$

Mais la valeur de H_0 n'est pas connue avec une précision absolue. Il règne sur cette quantité une incertitude :

$$\Delta H_0 = \frac{e}{c}\left[\frac{\Delta v_x}{y^2} + \frac{v_x}{y^3}\Delta y\right] \tag{4}$$

puisqu'il y a nécessairement une incertitude Δv_x sur la composante x de la vitesse de l'électron et une incertitude Δy sur son ordonnée y [1]. Mais d'après les relations d'Heisenberg, en nous bornant à l'approximation Newtonienne, nous avons :

$$\Delta v_x \geqslant \frac{h}{m_0\,\Delta x} \qquad \Delta v_y \geqslant \frac{h}{m_0\,\Delta y} \tag{5}$$

[1] Dans la formule (4), il faut le signe + dans le crochet parce que dans le cas le plus défavorable, les deux incertitudes peuvent ajouter leurs effets.

et d'autre part nous devons toujours supposer :

$$\Delta v_y \ll v_x \tag{6}$$

sans quoi nous n'aurions pas un mouvement s'opérant sensiblement le long de l'axe des x.

Si maintenant nous comparons (3) et (4), nous trouvons :

$$\frac{|\Delta H_0|}{H_P} = \frac{4 \pi m_0}{h} (y \, \Delta v_x + v_x \, \Delta y) \tag{7}$$

d'où résulte par (5), (1) et (6) :

$$\frac{|\Delta H_0|}{H_P} > \left(\frac{y}{\Delta x} + \frac{v_x}{\Delta v_y} \right) \gg 1 \tag{8}$$

L'incertitude sur le champ magnétique orbital est donc toujours beaucoup plus grande que la valeur du champ propre et par suite le magnétomètre ne nous permet pas de mesurer le moment magnétique propre de l'électron.

Reprenons le même genre de raisonnement pour l'action sur l'électron-aimant d'un champ magnétique non uniforme. Soit encore un électron se déplaçant le long de l'axe des x ; imaginons un champ magnétique parallèle à l'axe des y et possédant un gradient $\frac{\partial H}{\partial y}$ notable.

A l'origine des coordonnées, le champ magnétique a la valeur $H(0)$. En passant près de l'origine, l'électron en mouvement subit la force électrodynamique de Lorentz due au mouvement orbital ; cette force est sensiblement égale à :

$$f_0 = e \frac{v_x}{c} . H(0) \tag{9}$$

D'autre part l'aimant auquel l'électron est équivalent subit la force (due au moment magnétique propre) :

$$f_P = \frac{e h}{4 \pi m_0 c} \left(\frac{\partial H}{\partial y} \right)_{y=0} \tag{10}$$

Mais comme le train d'ondes a une largeur Δy et que l'ordonnée de l'électron est par suite affectée de cette incertitude, on a sur la valeur de f_0 l'incertitude :

$$|\Delta f_0| = \frac{e}{c}\left| v_x \left(\frac{\partial H}{\partial y}\right)_0 \Delta y + H(0)\,\Delta v_x \right| \tag{11}$$

d'où en comparant avec (10) :

$$\frac{|\Delta f_0|}{f_P} = \frac{4\pi m_0}{h}\left| v_x\,\Delta y + \left(\frac{H}{\frac{\partial H}{\partial y}}\right)_0 \Delta v_x \right| \tag{12}$$

Comme les relations (5) et (6) sont encore ici valables, on a donc :

$$\frac{|\Delta f_0|}{f_P} \geqslant \left| \frac{v_x}{\Delta v_y} + \left(\frac{H}{\frac{\partial H}{\partial y}}\right)_0 \cdot \frac{1}{\Delta x} \right| \geqslant 1. \tag{13}$$

L'action du champ magnétique extérieur non uniforme sur l'aimant électronique est donc complètement masquée par l'incertitude sur la force de Lorentz et cette méthode ne permet pas non plus la mesure du moment magnétique propre.

Ces raisonnements qu'on peut d'ailleurs perfectionner [1] rendent donc très vraisemblable l'impossibilité de mesurer directement le moment magnétique propre de l'électron.

2. ***Impossibilité de mesurer le moment de rotation propre.*** — On peut étendre, comme l'a en particulier signalé M. Darwin, le même genre de considérations au moment de rotation propre.

Imaginons par exemple une ouverture rectangulaire de côtés Δx et Δy percée dans un écran plan.

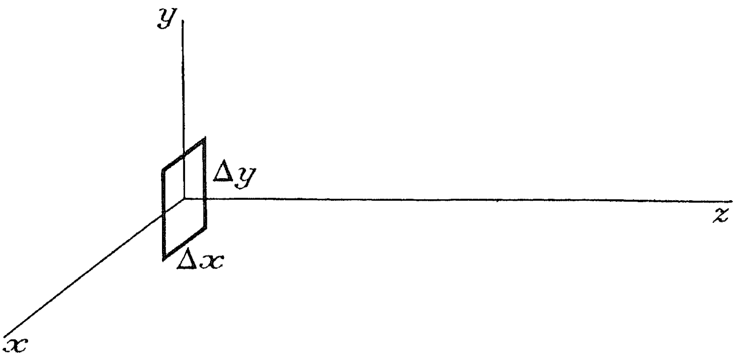
Sur l'écran, faisons tomber du côté gauche un électron associé à une onde plane monochromatique; le train d'ondes de l'électron à droite de l'écran a pour dimensions latérales Δx et Δy. Les composantes x et y de la quantité de mouve-

(1) Voir Darwin : *Examples of the Uncertainty Principle* (*Proc. of the Royal Society*, A. Vol. 130 [1931], p. 637) et le rapport de M. Pauli au Congrès Solvay de 1930 (Gauthier-Villars, éditeur).

ment de l'électron après le passage de l'écran sont donc affectées des incertitudes :

$$|\Delta p_x| \geqslant \frac{h}{\Delta x} \; ; \quad |\Delta p_y| \geqslant \frac{h}{\Delta y} \tag{14}$$

Or la composante z du moment orbital de l'électron, c'est-



à-dire du moment de rotation dû à son mouvement de translation, est :

$$M_z = x\,p_y - y\,p_x \tag{15}$$

Il a donc une valeur comprise entre 0 et :

$$\Delta M_z = \Delta x\,|\Delta p_y| + \Delta y\,|\Delta p_x| \tag{16}$$

d'où d'après (14) :

$$\Delta M_z \geqslant h\left(\frac{\Delta x}{\Delta y} + \frac{\Delta y}{\Delta x}\right) = h \cdot \frac{\Delta x^2 + \Delta y^2}{\Delta x\,\Delta y} \tag{17}$$

La fraction qui figure au dernier membre de (17) étant supérieure à 1, on en déduit *a fortiori* :

$$\Delta M_z > \frac{h}{4\pi} \tag{18}$$

L'incertitude sur la composante z du moment de rotation orbital est donc supérieure au moment de rotation propre et ceci rend illusoire la mesure de cette dernière grandeur.

Le même raisonnement peut s'appliquer aux composantes x et y du moment de rotation propre.

Ces considérations semblent établir que le moment de rotation propre de l'électron n'est pas plus mesurable expérimentalement que son mouvement magnétique propre.

3. ***Théorie plus générale.*** — On peut rattacher les considérations qui précèdent à un point de vue plus général comme l'a montré en particulier M. Pauli [1]. Pour le comprendre, nous allons rappeler très brièvement la méthode d'approximation dite « méthode de Brillouin-Wentzel ».

En Mécanique ondulatoire non-relativiste, le principe de la méthode de Brillouin-Wentzel [2] consiste à poser :

$$\Psi = e^{\frac{2\pi i}{h} S} \tag{19}$$

avec :

$$S = S_0 + \frac{h}{2\pi i} S_1 + \left(\frac{h}{2\pi i}\right)^2 S_2 + \cdots + \left(\frac{h}{2\pi i}\right)^n S_n + \cdots \tag{20}$$

et à déterminer par approximations successives les termes S_0, S_1... qui figurent dans le développement de S suivant les puissances successives de la quantité de très petit module $\frac{h}{2\pi i}$. Si dans ce développement on se contente des termes d'ordre zéro, c'est-à-dire de ceux qui subsistent quand on fait $h = 0$, on retrouve l'optique géométrique pour les ondes Ψ et la mécanique classique pour le corpuscule associé car, en ce cas, on peut considérer un groupe d'ondes de dimensions très petites décrivant un des rayons de l'onde et l'assimiler au corpuscule ponctuel de l'ancienne Mécanique : il y a coïncidence entre les rayons de l'onde et les trajectoires classiques. Mais si l'on tient compte des termes d'ordres 1, 2... en $\frac{h}{2\pi i}$ dans le développement (20), on voit

(1) *Helvetica Physica Acta*, vol. V, fasc. III, p. 179.

(2) Cf. L. Brillouin, *Exposés sur la théorie des Quanta*, fasc. I, Hermann éd.

apparaître les particularités qui opposent l'optique ondulatoire à l'optique géométrique, la nouvelle Mécanique à l'ancienne.

On peut reprendre la même méthode d'approximation en théorie de Dirac : c'est ce qu'a fait M. Pauli. Ici l'on pose :

$$\Psi_k = e^{\frac{2\pi i}{h} S_k} \qquad (k = 1, 2, 3, 4) \qquad (21)$$

avec :

$$S_k = S_{o,k} + \frac{h}{2\pi i} S_{1,k} + \cdots + \left(\frac{h}{2\pi i}\right)^n S_{n,k} + \cdots \qquad (22)$$

et l'on déterminera les $S_{\nu k}$ par approximations successives. En procédant ainsi, il apparaît que, si l'on conserve seulement les termes d'ordre 0, on obtient l'optique géométrique qui correspond à la Mécanique ondulatoire *relativiste sans spin*, c'est-à-dire la Mécanique ancienne sous la forme Einsteinienne. On peut alors imaginer des trains d'ondes extrêmement petits décrivant des rayons-trajectoire qui coïncident avec les trajectoires prévues par la Dynamique Einsteinienne de l'électron. Or la Mécanique Einsteinienne considère l'électron comme un simple corpuscule chargé et ignore le « spin ». Donc à l'approximation d'ordre zéro, la Mécanique de Dirac nous ramène à la dynamique relativiste de l'électron « sans spin ». Ceci pouvait être prévu puisque le moment magnétique propre et le moment de rotation propre de l'électron sont proportionnels à h et disparaissent si l'on néglige les termes de l'ordre de h.

Si, poursuivant l'approximation, on tient compte dans l'équation (22) des termes d'ordre 1, 2 ... en $\frac{h}{2\pi i}$, on voit apparaître les termes qui traduisent l'existence du magnétisme et de la rotation propres de l'électron, mais en même temps, comme dans la théorie non relativiste, on sort du domaine d'application de l'optique géométrique et en conséquence les concepts de l'ancienne Mécanique ponctuelle cessent d'être exacts.

On comprend maintenant pourquoi, conformément aux

conclusions de M. Bohr, une expérience dans laquelle il est possible de traiter l'électron comme un point matériel ne peut pas conduire à mettre en évidence le moment magnétique propre ou la rotation propre de l'électron. Il serait en effet, d'après l'analyse de M. Pauli, contradictoire de supposer que l'ancienne Mécanique ponctuelle est valable et que les effets caractéristiques du « spin » peuvent se manifester.

4. ***Polarisation des ondes électroniques.*** — S'il paraît impossible de mesurer le moment magnétique d'un électron considéré comme un corpuscule aimanté, rien ne s'oppose *a priori* à la mise en évidence expérimentale du caractère de polarisation qu'impose à l'onde Ψ électronique l'existence de ce moment magnétique. Cette polarisation des ondes Ψ diffère d'ailleurs notablement de la polarisation classique des ondes lumineuses. Tandis que cette dernière est définie pour une onde plane par un vecteur oscillant toujours normal à la direction de propagation, la polarisation de l'onde Ψ de Dirac est définie par le vecteur I « densité de moment magnétique » et ce vecteur dans une onde plane est, comme nous le savons, constant et orienté n'importe comment par rapport à la direction de la propagation. Par suite, tandis que les propriétés d'un faisceau de lumière polarisée examinées dans les divers azimuts autour de sa direction de propagation, présentent toujours la période π pour un faisceau d'ondes Ψ de Dirac la période correspondante est 2π.

Pour mettre en évidence la polarisation des ondes électroniques, on peut songer à employer des dispositifs analogues à l'appareil de Nörremberg en optique. Supposons un faisceau d'électrons non polarisé, c'est-à-dire un faisceau dans lequel les vecteurs $\vec{I}$ relatifs aux divers électrons sont orientés au hasard. Si l'on fait se réfléchir ce faisceau sur un corps cristallisé, on conçoit par analogie avec l'optique que le faisceau réfléchi puisse être partiellement polarisé. Si l'on fait tomber ce faisceau réfléchi sur un deuxième réflecteur, la deuxième

réflexion s'effectuera avec plus ou moins d'intensité suivant l'azimut du plan d'incidence. La théorie exacte du phénomène paraît assez compliquée et n'a pas été encore complètement développée ; nous ne l'aborderons pas ici. En gros, on est conduit à prévoir un effet très faible pour les électrons de quelques milliers de volts, effet qui doit croître avec l'énergie. Au point de vue expérimental, le phénomène ne paraît pas avoir été observé avec certitude. M. Rupp a bien publié des photographies sur lesquelles on voit nettement une influence de l'azimut sur la seconde réflexion, mais ces résultats ne paraissent pas confirmés jusqu'à présent par ceux des autres expérimentateurs. La question reste ouverte (1).

(1) Les principaux travaux théoriques sur la polarisation des électrons sont ceux de M. Mott (*Proc. Roy. Soc.*, A 124 [1929], p. 425 et A. 135 [1932], p. 429). On trouvera une bibliographie complète dans un mémoire de MM. Thibaud, Trillat et v. Hirsch (*J. Phys.*, VII, t. 33 [1932], p. 314).

CHAPITRE XX

Les états d'énergie négative en théorie de Dirac

1. *L'onde plane à énergie négative*

NOUS allons aborder maintenant une des grosses difficultés soulevées par la théorie de Dirac.

Nous avons précédemment étudié la forme des fonctions Ψ_k pour une onde plane monochromatique dans le cas de l'absence de champ (voir chapitre XII, parag. 3). Pour cela, nous avons écrit l'équation de Dirac avec potentiels nuls et nous avons essayé une solution de la forme :

$$\Psi_k = a_k \, e^{\frac{2\pi i}{h}(Wt - p_x x - p_y y - p_z z)} \tag{1}$$

Nous avons ainsi trouvé pour les a_k les équations linéaires et homogènes :

$$\begin{aligned}
&\left(\frac{W}{c} + m_0 c\right) a_1 + (p_x + i\, p_y)\, a_4 + p_z\, a_3 = 0 \\
&\left(\frac{W}{c} + m_0 c\right) a_2 + (p_x - i\, p_y)\, a_3 - p_z\, a_4 = 0 \\
&\left(\frac{W}{c} - m_0 c\right) a_3 + (p_x + i\, p_y)\, a_2 + p_z\, a_1 = 0 \\
&\left(\frac{W}{c} - m_0 c\right) a_4 + (p_x - i\, p_y)\, a_1 - p_z\, a_2 = 0
\end{aligned} \tag{2}$$

Pour qu'il y ait une solution non nulle, il faut que le déterminant des équations (2) soit nul. Ceci nous a donné la condition :

$$\frac{W^2}{c^2} = m_0^2c^2 + p_x^2 + p_y^2 + p_z^2 \tag{3}$$

ce qui est la relation relativiste classique. En posant alors :

$$W = + c\sqrt{m_0^2c^2 + p_x^2 + p_y^2 + p_z^2} \tag{4}$$

nous avons trouvé la solution :

$$a_1 = -\frac{p_zA + (p_x + ip_y)B}{\frac{W}{c} + m_0c}; \quad a_2 = -\frac{(p_x - ip_y)A - p_zB}{\frac{W}{c} + m_0c}$$

$$a_3 = A; \quad a_4 = B \tag{5}$$

où A et B sont des constantes complexes arbitraires.

Mais nous aurions pu aussi satisfaire à la condition (3) en posant :

$$W = - c\sqrt{m_0^2c^2 + p_x^2 + p_y^2 + p_z^2}. \tag{6}$$

Nous aurions alors trouvé la solution suivante :

$$a_1 = C; \ a_2 = D; \ a_3 = \frac{p_zC + (p_x + ip_y)D}{m_0c - \frac{W}{c}};$$

$$a_4 = \frac{(p_x - ip_y)C - p_zD}{m_0c - \frac{W}{c}}. \tag{7}$$

C et D étant deux constantes complexes arbitraires.

Nous allons maintenant modifier légèrement les notations que nous venons d'employer. Pour des valeurs données de p_x, p_y et p_z, nous définirons toujours à l'avenir W par la formule (4) avec le signe + et pour tenir compte de la solution (7), nous dirons que nous avons à considérer à la fois l'onde d'énergie + W et l'onde d'énergie — W. Avec cette nouvelle convention, il faut dans les équations (7) changer W en — W.

Bref pour une valeur donnée de p_x, p_y et p_z, W étant défini par la relation (4), nous avons à considérer l'onde plane monochromatique à énergie positive $+W$ définie par :

$$\Psi_k = a_k \, e^{\frac{2\pi i}{h}(Wt - p_x x - p_y y - p_z z)} \tag{8}$$

avec :

$$a_1 = -\frac{p_z A + (p_x + ip_y) B}{\frac{W}{c} + m_0 c} \; ; \quad a_2 = -\frac{(p_x - ip_y) A - p_z B}{\frac{W}{c} + m_0 c} \; ;$$

$$a_3 = A \; ; \quad a_4 = B \tag{9}$$

et l'onde plane monochromatique à énergie négative $-W$ définie par :

$$\Psi_k = b_k \, e^{\frac{2\pi i}{h}(-Wt - p_x x - p_y y - p_z z)} \tag{10}$$

avec :

$$b_1 = C \; ; \quad b_2 = D \; ; \quad b_3 = \frac{p_z C + (p_x + ip_y) D}{\frac{W}{c} + m_0 c} \; ;$$

$$b_4 = \frac{(p_x - ip_y) C - p_z D}{\frac{W}{c} + m_0 c} . \tag{11}$$

Maintenant, examinons ces deux ondes.

Pour l'onde à énergie positive, nous savons déjà que les composantes Ψ_3 et Ψ_4, qui correspondent en quelque sorte à la masse propre positive $+m_0$, l'emportent sur les ondes Ψ_1 et Ψ_2 qui correspondent à la masse propre négative $-m_0$; ces ondes Ψ_1 et Ψ_2, nulles si la vitesse est nulle, ne prennent de l'importance que pour les vitesses suffisamment voisines de celles de la lumière. Ce serait seulement dans le cas limite $v = c$ que l'on aurait :

$$|\Psi_1|^2 + |\Psi_2|^2 = |\Psi_3|^2 + |\Psi_4|^2 .$$

Autrement dit, on a toujours :

$$\Omega_1 = |\Psi_1|^2 + |\Psi_2|^2 - |\Psi_3|^2 - |\Psi_4|^2 \leqslant 0 \tag{12}$$

l'égalité se rapportant au cas limite $v = c$.

Pour l'onde à énergie négative, les conclusions sont tout l'opposé. Les ondes Ψ_1 et Ψ_2 sont ici prépondérantes. Si l'électron est en repos, c'est-à-dire si p_x, p_y et p_z sont nuls, on a :

$$\Psi_1 = \mathrm{C}\, e^{-\frac{2\pi i}{h} m_0 c^2 t}\ ; \quad \Psi_2 = \mathrm{D}\, e^{-\frac{2\pi i}{h} m_0 c^2 t}\ ; \quad \Psi_3 = \Psi_4 = 0. \qquad (13)$$

Pour des vitesses croissantes, les ondes Ψ_3 et Ψ_4 prennent de l'importance, mais ce serait seulement pour le cas limite $v = c$ que l'on aurait l'égalité entre

$$|\Psi_1|^2 + |\Psi_2|^2 \text{ et } |\Psi_3|^2 + |\Psi_4|^2.$$

Donc, on a ici :

$$\Omega_1 \geqslant 0 \qquad (14)$$

l'égalité se rapportant au cas limite $v = c$. Ici, ce sont en quelque sorte les ondes correspondant à une masse propre négative qui l'emportent.

L'existence d'états à énergie négative dans la théorie de Dirac constitue une grave difficulté pour cette théorie parce qu'un électron qui serait dans un tel état aurait des propriétés tout à fait étranges qui n'ont jamais été observées. Placé dans un champ électrique $\vec{h}$, il prendrait une accélération de sens opposé à la force $-e\vec{h}$; on augmenterait sa vitesse en lui retirant de l'énergie ; sa vitesse serait en sens inverse de sa quantité de mouvement [1] ; etc. La théorie de Dirac devrait donc, semble-t-il, s'efforcer d'éliminer ces états à énergie négative puisqu'ils ne semblent pas répondre à une réalité ; mais pour des raisons que nous allons maintenant signaler, cette élimination ne paraît pas facile.

2. ***Caractère incomplet du système des ondes à énergie positive.*** — Revenons sur certaines particularités de la Mécanique ondulatoire non relativiste.

---

(1) En vertu de la relation $v = \frac{\partial W}{\partial p}$ qui exprime l'égalité de la vitesse du corpuscule et de la vitesse de groupe des ondes associées (Cf. *Introd. à l'étude de la Mécanique ondulatoire*, p. 75).

En Mécanique ondulatoire non relativiste, l'équation de propagation :

$$\frac{h}{2 \pi i} \frac{\partial \Psi}{\partial t} = \mathrm{H}(\Psi) \tag{15}$$

est du premier ordre par rapport au temps. La solution de cette équation est donc complètement déterminée si l'on connaît sa forme initiale $\Psi(x, y, z, 0)$.

Considérons le cas d'absence de champ. L'équation (15) s'écrit alors :

$$\Delta \Psi = \frac{4 \pi i m}{h} \frac{\partial \Psi}{\partial t} \tag{16}$$

et admet pour solution l'onde plane monochromatique :

$$\Psi(x, y, z, t) = a \, e^{\frac{2 \pi i}{h}(\mathrm{E} t - p_x x - p_y y - p_z z)} \tag{17}$$

avec (sans ambiguïté de signe) :

$$\mathrm{E} = \frac{1}{2 m}\left[p_x^2 + p_y^2 + p_z^2\right] \tag{18}$$

Supposons que nous nous donnions la forme initiale de la fonction d'ondes :

$$\Psi(x, y, z, 0) = \mathrm{F}(x, y, z) \tag{19}$$

Supposons que $\mathrm{F}(x, y, z)$ soit développable en intégrale de Fourier sous la forme :

$$\mathrm{F}(x, y, z) = \int\!\!\int\!\!\int_{-\infty}^{+\infty} g(p_x, p_y, p_z) \, e^{-\frac{2 \pi i}{h}(p_x x + p_y y + p_z z)} dp_x \, dp_y \, dp_z \, . \tag{20}$$

Les coefficients de ce développement sont calculables, à partir de la fonction F donnée, par la formule :

$$g(p_x, p_y, p_z) = \frac{1}{h^3} \int\!\!\int\!\!\int_{-\infty}^{+\infty} \mathrm{F}(x, y, z) \, e^{\frac{2 \pi i}{h}(p_x x + p_y y + p_z z)} dx \, dy \, dz \, . \tag{21}$$

Je dis alors que la fonction :

$$\Psi(x, y, z, t) = \int\int\int_{-\infty}^{+\infty} g(p_x, p_y, p_z)\, e^{\frac{2\pi i}{h}(Et - p_x x - p_y y - p_z z)}\, dp_x\, dp_y\, dp_z \quad (22)$$

Ceci est évident puisque : 1° Ψ est solution de l'équation (16) comme étant une somme de solutions planes monochromatiques de cette équation linéaire ; 2° pour $t=0$ on a bien $\Psi(x, y, z, 0) = F(x, y, z)$.

Nous en concluons le théorème suivant : « Dans la Mécanique ondulatoire non relativiste du point matériel libre, les ondes planes monochromatiques constituent un système « complet », c'est-à-dire qu'il est possible de représenter n'importe quelle solution par une superposition de telles ondes. »

Passons maintenant à la théorie de Dirac et cherchons à y transposer les raisonnements que nous venons de faire. Ici, nous avons 4 équations simultanées du premier ordre entre les 4 Ψ_k. Ces 4 fonctions d'onde sont donc entièrement déterminées si l'on se donne les formes initiales $\Psi_k(x, y, z, 0)$.

Prenons toujours le cas du champ nul et demandons-nous encore s'il est possible de représenter n'importe quelle solution par une superposition d'ondes planes monochromatiques. Une solution quelconque est entièrement définie par les 4 fonctions :

$$\Psi_k(x, y, z, 0) = F_k(x, y, z) \qquad (k = 1, 2, 3, 4) \qquad (23)$$

Supposons les F_k arbitrairement données et développables par le théorème de Fourier sous la forme :

$$F_k(x, y, z) = \int\int\int_{-\infty}^{+\infty} g_k(p_x, p_y, p_z)\, e^{-\frac{2\pi i}{h}(p_x x + p_y y + p_z z)}\, dp_x\, dp_y\, dp_z \quad (24)$$

les g_k étant donnés par :

$$g_k(p_x, p_y, p_z) = \frac{1}{h^3}\int\int\int_{-\infty}^{+\infty} F_k(x, y, z)\, e^{\frac{2\pi i}{h}(p_x x + p_y y + p_z z)}\, dx\, dy\, dz\,. \quad (25)$$

Cherchons à représenter la solution qui correspond aux Ψ_k initiaux donnés par une superposition d'ondes planes monochromatiques qui ne contiendrait que des ondes à énergie positive. Pour cela posons :

$$\Psi_k(x, y, z, t) = \int\int\int_{-\infty}^{+\infty} a_k(p_x, p_y, p_z)\, e^{\frac{2\pi i}{h}(Wt - p_x x - p_y y - p_z z)}\, dp_x\, dp_y\, dp_z. \tag{26}$$

W étant défini par la relation (4). Ces fonctions (26) donneraient bien une solution de l'équation de propagation, mais pour qu'elles prennent les formes intiales $F_k(x, y, z)$, il faudrait avoir :

$$a_k(p_x, p_y, p_z) = g_k(p_x, p_y, p_z) \qquad (k = 1, 2, 3, 4) \tag{27}$$

les g_k étant connus. Mais nous savons que, sur les quatre $a_k(p_x, p_y, p_z)$, deux seulement peuvent être choisis arbitrairement et ceci nous montre que l'on ne peut pas en général satisfaire aux conditions (27). Les ondes planes monochromatiques à énergie positive ne forment donc pas un système complet pour l'électron de Dirac en l'absence de champ.

Au contraire, en considérant, à côté des ondes planes à énergie positive, aussi les ondes planes à énergie négative, nous obtenons un système complet. En effet, si nous posons:

$$\begin{aligned}\Psi_k(x, y, z, t) = &\int\int\int_{-\infty}^{+\infty} a_k(p_x, p_y, p_z)\, e^{\frac{2\pi i}{h}(Wt - p_x x - p_y y - p_z z)}\, dp_x\, dp_y\, dp_z \\ &+ \int\int\int_{-\infty}^{+\infty} b_k(p_x, p_y, p_z)\, e^{\frac{2\pi i}{h}(-Wt - p_x x - p_y y - p_z z)}\, dp_x\, dp_y\, dp_z\end{aligned} \tag{28}$$

W étant toujours défini par (4), nous aurons une solution des équations de propagation puisque celles-ci sont linéaires et pour que les Ψ_k initiaux coïncident avec les $F_k(x, y, z)$ données, nous aurons à écrire la condition :

$$a_k(p_x, p_y, p_z) + b_k(p_x, p_y, p_z) = g_k(p_x, p_y, p_z); \; (k = 1, 2, 3, 4). \tag{29}$$

Or, à la différence des conditions (27), les conditions (29) sont compatibles car, sur les 8 a_k et b_k qui correspondent à un ensemble $p_x p_y p_z$, il y en a 4 d'arbitraires.

Si nous écrivons explicitement les conditions (29), nous obtenons pour chaque ensemble p_x, p_y, p_z, les conditions :

$$
\begin{aligned}
&-\frac{(p_x + ip_y)\,\mathrm{A} + p_z\,\mathrm{B}}{\frac{\mathrm{W}}{c} + m_0 c} + \mathrm{C} = g_1\,(p_x, p_y, p_z)\,; \\
&-\frac{(p_x - ip_y)\,\mathrm{A} - p_z\,\mathrm{B}}{\frac{\mathrm{W}}{c} + m_0 c} + \mathrm{D} = g_2\,(p_x\, p_y\, p_z) \\
&\frac{p_z \mathrm{C} + (p_x + ip_y)\,\mathrm{D}}{\frac{\mathrm{W}}{c} + m_0 c} + \mathrm{A} = g_3\,(p_x\, p_y\, p_z)\,; \\
&\frac{(p_x - ip_y)\,\mathrm{C} + p_z\,\mathrm{D}}{\frac{\mathrm{W}}{c} + m_0 c} + \mathrm{B} = g_4\,(p_x\, p_y\, p_z).
\end{aligned}
\qquad (30)
$$

Si l'on étudie le système (30) en tenant compte des relations d'Heisenberg, on voit que, dans les cas les plus favorables, il est possible de représenter un train d'ondes Ψ par une superposition d'ondes à énergie positive mais seulement si les dimensions du train d'ondes sont sensiblement supérieures à $\frac{h}{m_0 c}$. Par contre, si les dimensions du train d'ondes sont inférieures à $\frac{h}{m_0 c}$, il est en général tout à fait impossible de le représenter par une superposition d'ondes planes sans faire appel à des ondes à énergie négative.

D'une façon générale, il est donc impossible en théorie de Dirac de représenter un train d'ondes quelconque sans faire intervenir les ondes à énergie négative et cette impossibilité nous prouve combien il est difficile d'arriver à éliminer ces ondes.

3. ***Le paradoxe de Klein.*** — Au premier abord, on pour-

rait croire que la difficulté des énergies négatives existe même en Relativité classique. En effet, dans la Relativité classique, l'énergie est définie en fonction de la quantité de mouvement dans le cas de l'absence de champ par la relation (3), relation qui donne pour W les deux valeurs :

$$W = \pm c\sqrt{m_0^2 c^2 + p_x^2 + p_y^2 + p_z^2} \tag{31}$$

Mais il faut remarquer que, d'après (31), les valeurs possibles de W sont comprises dans deux domaines séparés de $+\infty$ à $+m_0 c^2$ et de $-m_0 c^2$ à $-\infty$. L'intervalle de $+m_0 c^2$ à $-m_0 c^2$ ne correspond à aucune valeur possible de l'énergie. Or en Dynamique ancienne, même relativiste, les grandeurs mécaniques et en particulier l'énergie varient en principe d'une façon continue. Si donc à l'origine les électrons avaient leurs énergies comprises dans le domaine positif $m_0 c^2 \rightarrow +\infty$, il en sera toujours de même par la suite et aucune valeur de l'énergie comprise dans le domaine négatif $-m_0 c^2 \rightarrow -\infty$ ne pourra apparaître puisque les deux domaines sont séparés par l'intervalle $+m_0 c^2 \rightarrow -m_0 c^2$ que les valeurs de l'énergie ne peuvent pas traverser. L'objection tirée de l'existence des énergies négatives se trouve ainsi écartée en Dynamique Einsteinienne.

Il n'en est pas de même en Mécanique nouvelle car celle-ci admet en principe la possibilité des transitions brusques entre états dont l'énergie diffère d'une quantité finie, ce qui empêche d'écarter *a priori* le passage du domaine des énergies positives à celui des énergies négatives. Et, qui plus est, il est facile d'imaginer des exemples simples où des transitions de ce genre se trouvent réalisées.

M. O. Klein a le premier (¹) signalé un exemple de transition qui, sans être à proprement parler un passage d'un état d'énergie positive à un état d'énergie négative, lui est cependant équivalente.

M. Klein considère une surface plane S qui sépare une

(¹) *Zeitschrift für Physik*, 53 (1929), p. 157.

région I où les potentiels sont nuls d'une région II où règne un potentiel scalaire V constant et négatif de sorte que dans la région II un électron a une énergie potentielle $U = -eV > 0$. Sur la surface de séparation, tombe normalement venant de la région I une onde électronique de Dirac; cette onde est supposée monochromatique plane et correspondant à l'énergie positive W. Il s'agit de calculer les ondes réfléchies et transmises par la surface de séparation. On montre que pour faire ce calcul, il faut d'abord exprimer qu'il y a continuité de chacun des 4 Ψ_k au travers de la surface de séparation, c'est-à-dire écrire les 4 équations :

$$\Psi_k \text{ incident} + \Psi_k \text{ réfléchi} = \Psi_k \text{ transmis} \qquad (32)$$

Naturellement les ondes réfléchies et transmises correspondent à la même énergie que l'onde incidente : le phénomène est conservatif.

Ceci posé, M. Klein démontre les résultats suivants que je me contente d'énoncer sans démonstration. Pour $0 < U < W - m_0 c^2$, il y a à la fois réflexion et transmission, l'onde transmise ayant comme l'onde réfléchie les caractères normaux d'une onde à énergie positive.

Pour $W - m_0 c^2 < U < W + m_0 c^2$, il y a alors réflexion totale avec onde évanescente dans le second milieu. Pour $U > W + m_0 c^2$, on trouve à nouveau une onde transmise à travers la surface, mais, et c'est là le résultat fondamental, cette onde est une sorte d'onde à énergie négative ; évidemment elle correspond à l'énergie totale W qui est positive mais ce qu'on pourrait appeler « l'énergie de nature non potentielle » de l'électron dans la région II, c'est-à-dire la quantité $W - U$, est négative et inférieure à $-m_0 c^2$ alors que dans la Dynamique Einsteinienne classique cette quantité est toujours supérieure à $m_0 c^2$. L'onde transmise dans le milieu II où règne le potentiel scalaire V est l'analogue d'une onde à énergie négative en l'absence de potentiel et possède les mêmes propriétés paradoxales. L'existence de cette onde transmise doit être interprétée en disant qu'il y a une cer-

taine probabilité pour qu'un électron incident pénètre dans la région II en passant dans cet état étrange.

D'après Klein, cette probabilité peut même être notable. Il semble bien que l'on ne puisse considérer le résultat du calcul comme physiquement exact et c'est là le paradoxe de Klein.

Il est vrai que le cas considéré par M. Klein est extrêmement schématique. D'autres auteurs ont envisagé d'autres exemples un peu moins artificiels. Le résultat général de ces recherches paraît être le suivant : chaque fois que l'énergie potentielle de l'électron doit subir une variation d'au moins $m_0 c^2$ sur une distance inférieure à $\frac{h}{m_0 c}$, il y a possibilité d'apparition des états d'énergie négative. Ce résultat conduit à penser que si l'on pouvait s'interdire de considérer des distances spatiales inférieure à $\frac{h}{m_0 c}$, on parviendrait peut-être à éliminer les ondes à énergie négative. Ceci doit être rapproché de ce qui a été dit à la fin du paragraphe précédent au sujet de la représentation des trains d'ondes.

4. ***Remarques et conclusions.*** — Les états d'énergie négative interviennent encore d'une façon curieuse dans la théorie de la diffusion de la lumière par l'électron de Dirac. Nous ne développerons pas cette théorie en renvoyant le lecteur à d'autres exposés (1). Nous en noterons seulement le résultat : l'électron de Dirac ne pourrait pas diffuser la lumière s'il n'était pas susceptible d'états d'énergie négative. Comme il est nécessaire d'admettre que les électrons diffusent la lumière pour expliquer les phénomènes de diffusion par les corps matériels, ceci montre encore combien il est difficile d'affranchir la théorie de Dirac de cette apparente im-

(1) Voir notamment le remarquable exposé de E. FERMI : *Quantum theory of radiation* (*Reviews of Modern Physics*, vol. 4 [1932], p. 120).

perfection qu'est l'existence des états d'énergie négative. Diverses tentatives ont été faites néanmoins pour tâcher de tourner la difficulté : nous en dirons seulement quelques mots.

M. Schrödinger a proposé une modification très ingénieuse des équations générales de Dirac qui ferait disparaître les états d'énergie négative (1). Mais outre que cette modification paraît difficilement compatible avec l'existence de la diffusion de la lumière par les électrons, elle a un caractère assez nettement artificiel.

M. Dirac, au lieu de vouloir supprimer les états d'énergie négative, a au contraire cherché à les interpréter (2). Il a pour cela supposé que ces états existent réellement et que dans toute partie de l'espace se trouverait un nombre infini d'électrons occupant tous ces états d'énergie négative, électrons qui seraient inobservables. Quelques-uns de ces électrons quitteraient de temps à autre leur état habituel d'énergie négative pour prendre un état d'énergie positive et ceux-là seraient les électrons observables. Le « trou » laissé dans les états à énergie négative par le départ d'un électron serait ce que nous appelons un proton. Le retour de l'électron évadé dans un état d'énergie négative constituerait la disparition simultanée d'un électron et d'un proton, disparition qui doit s'accompagner d'une émission de radiation. Malheureusement ces séduisantes hypothèses se sont heurtées à toutes sortes d'objections et ne semblent pas pouvoir être conservées (1).

En résumé, les états d'énergie négative de l'électron jouent un rôle très important dans la structure même de la théorie de Dirac et cependant ils ne se manifestent pas dans la réalité. Il semble que, pour l'instant, on ne puisse que constater la difficulté sans la résoudre.

---

(1) Voir notamment *Annales de l'Institut Henri Poincaré*, t. II, p. 269.

(2) *Ibidem*, t. I, p. 357.

(3) La théorie des trous de Dirac connaît en ce moment un retour de faveur par suite de la découverte expérimentale de l'électron positif.

CHAPITRE XXI

Le tremblement de Schrödinger

1. *Le mouvement du centre de gravité de la probabilité.*

NOUS éviterons dans ce chapitre d'introduire explicitement un opérateur correspondant à la « vitesse » de l'électron. La vitesse d'un corpuscule est une notion dont on doit user prudemment dans la nouvelle Mécanique comme l'a remarqué M. Bohr. Elle n'est bien bien définie que dans certains cas et il ne paraît pas justifié de la considérer comme une grandeur physique observable.

Au contraire, il est toujours permis d'envisager la position moyenne du corpuscule ou centre de gravité de la probabilité et d'étudier son mouvement. Ce point est en effet défini en théorie de Dirac par ses coordonnées :

$$\bar{x} = \int\int\int x \sum_1^4{}_k \Psi_k^* \Psi_k \, d\tau \; ; \quad \bar{y} = \int\int_{-\infty}^{+\infty}\int y \sum_1^4{}_k \Psi_k^* \Psi_k \, d\tau \; ;$$

$$\bar{z} = \int\int_{-\infty}^{+\infty}\int z \sum_1^4{}_k \Psi_k^* \Psi_k \, d\tau \, . \qquad (1)$$

La vitesse du point de coordonnées $\bar{x}$, $\bar{y}$, $\bar{z}$ est, elle, une quantité parfaitement déterminée.

Soit $(\Psi_1 \Psi_2 \Psi_3 \Psi_4)$ la solution de l'équation de Dirac qui représente l'onde associée au mouvement d'un certain électron. Nous pouvons, nous le savons, développer chaque Ψ_k en série de fonctions propres sous la forme :

$$\Psi_k = \sum_n c_n \Psi_{k,n} \tag{2}$$

les c_n étant des constantes complexes. En substituant dans (1), on a :

$$\bar{x} = \sum_{m,n} c_m^* c_n \int_D \sum_1^4 {}_k (\Psi^*_{k,m} x \Psi_{k,n}) d\tau \tag{3}$$

en remplaçant par un seul signe de sommation les intégrales triples des formules (1). En désignant par x_{mn} l'élément d'indices m, n de la matrice correspondant à l'opérateur x, la formule (3) s'écrit simplement :

$$\bar{x} = \sum_{m,n} c_m^* c_n x_{mn} \tag{4}$$

Or, en vertu de la formule (25) du chapitre XV, on a :

$$\frac{d x_{mn}}{d t} = \int_D \sum_1^4 {}_k \left[\Psi^*_{k,m} \frac{2 \pi i}{h} (x \mathrm{H} - \mathrm{H} x) \Psi_{k,n}\right] d\tau \tag{5}$$

où H est l'opérateur Hamiltonien de Dirac :

$$\mathrm{H} = -[e \mathrm{V} + c(\alpha_1 \mathrm{P}_1 + \alpha_2 \mathrm{P}_2 + \alpha_3 \mathrm{P}_3 + \alpha_4 m_0 c)] \tag{6}$$

On trouve aisément :

$$\frac{2 \pi i}{h} \left(x \mathrm{H} - \mathrm{H} x\right) = c \alpha_1 \left(x \cdot \frac{\partial}{\partial x} - \frac{\partial}{\partial x} \cdot x\right) = - c \alpha_1 \tag{7}$$

et par suite :

$$\frac{d x_{mn}}{d t} = \int_{\bar{D}} \sum_1^4 {}_k [\Psi^*_{k,m} (- c \alpha_1) \Psi_{k,n}] d\tau \tag{8}$$

De (4), on obtient donc :

$$\frac{d\bar{x}}{dt} = \sum_{m,n} c_m^* c_n \frac{dx_{mn}}{dt}$$

$$= \int_D \sum_1^4{}_k \left[\sum_m c_m^* \Psi^*_{k,m} (-c\,\alpha_1) \sum_n c_n \Psi_{k,n} \right] d\tau$$

$$= \int_D \sum_1^4{}_k \Psi_k^* (-c\,\alpha_1) \Psi_k\, d\tau \qquad (9)$$

On trouverait naturellement de même :

$$\frac{d\bar{y}}{dt} = \int_D \sum_1^4{}_k \Psi_k^* (-c\,\alpha_2) \Psi_k d\tau; \quad \frac{d\bar{z}}{dt} = \int_D \sum_1^4{}_k \Psi_k^* (-c\,\alpha_3) \Psi_k \cdot d\tau. \qquad (10)$$

Nous avons trouvé précédemment (formules [7] du chap. XII) les expressions suivantes pour les composantes du courant de probabilité :

$$\rho\, u_x = -c \sum_1^4{}_k \Psi_k^* \alpha_1 \Psi_k\, d\tau. \qquad (11)$$

On a donc, par comparaison avec (9) :

$$\frac{d\bar{x}}{dt} = \int_D \rho\, u_x\, d\tau = \bar{u}_x \qquad (12)$$

et de même :

$$\frac{d\bar{y}}{dt} = \bar{u}_y \qquad \frac{d\bar{z}}{dt} = \bar{u}_z\,. \qquad (12')$$

La vitesse du centre de gravité de la probabilité est donc égale à la valeur moyenne de la vitesse de la probabilité, résultat qu'on aurait pu prévoir *a priori*.

On interprète souvent les formules (9) et (10) en disant que les opérateurs $-c\alpha_1$, $-c\alpha_2$ et $-c\alpha_3$ sont les opérateurs correspondant aux trois composantes de la vitesse de l'électron. Comme ces opérateurs n'ont pour valeurs propres que $+c$ et $-c$, on est amené à dire d'après les principes généraux de la nouvelle Mécanique que les seules valeurs pos-

sibles des composantes de la vitesse sont $+c$ et $-c$, résultat assez difficile à comprendre. Nous préférons, comme nous l'avons dit, nous abstenir de faire correspondre des opérateurs à la « vitesse » d'un corpuscule.

2. *Le théorème d'Ehrenfest n'est plus exact en théorie de Dirac.* — En Mécanique ondulatoire non relativiste, nous avons démontré le théorème d'Ehrenfest qui s'exprime par les formules :

$$m\frac{d^2\overline{x}}{dt^2}=\overline{f_x}\,;\quad m\frac{d^2\overline{y}}{dt^2}=\overline{f_y}\,;\quad m\frac{d^2\overline{z}}{dt^2}=\overline{f_z}\,. \tag{13}$$

Appliqué au cas de l'absence de champ, ce théorème nous conduit au résultat suivant : « En l'absence de champ, le mouvement du centre de gravité de la probabilité est rectiligne et uniforme. » Ceci est en quelque sorte la transposition en Mécanique ondulatoire du principe de l'inertie.

En théorie de Dirac, les résultats précédents ne sont plus exacts en général. En effet, partons de la formule (8) et appliquons encore la formule (25) du chapitre XV; nous obtenons :

$$\frac{d^2x_{m,n}}{dt^2}=\int_D\sum_1^4{}_k\left[\Psi_{k,m}{}^*\cdot\frac{2\pi i}{h}(-c\,\alpha_1\,H+H\,c\,\alpha_1)\cdot\Psi_{k,n}\right]\cdot d\tau \tag{14}$$

et par suite d'après (4) :

$$\frac{d^2\overline{x}}{dt^2}=\sum_{m,n}c_m^*\,c_n\int_D\sum_1^4{}_k\left[\Psi^*_{k,m}\frac{2\pi i}{h}(-c\alpha_1\,H+Hc\alpha_1)\,\Psi_{k,n}\right]\cdot d\tau. \tag{15}$$

Or α_1 ne commute pas avec H puisqu'il anticommute avec α_2, α_3 et α_4. On a donc en général

$$\frac{d^2\overline{x}}{dt^2}\neq 0\,,\ \text{et de même :}\ \frac{d^2\overline{y}}{dt^2}\neq 0\,;\ \frac{d^2\overline{z}}{dt^2}\neq 0\,. \tag{16}$$

Le mouvement du centre de gravité de la probabilité en l'absence de champ n'est pas en général rectiligne et uniforme.

Nous allons maintenant examiner de près la raison qui empêche ce mouvement du centre de gravité d'être rectiligne et uniforme et nous allons voir qu'elle est reliée à l'existence des états d'énergie négative. Nous obtiendrons ainsi sous une forme différente des résultats obtenus par M. Schrödinger (1).

3. ***Le tremblement de Schrödinger.*** — Pour bien comprendre pourquoi le centre de gravité de la probabilité n'a pas en général, même en l'absence de champ, un mouvement rectiligne et uniforme dans la Mécanique de Dirac, nous allons soumettre à une analyse détaillée l'expression :

$$\overline{\frac{dx}{dt}} = \int_D \sum_1^4{}_k \Psi_k^* (-c\,\alpha_1)\,\Psi_k\,d\tau. \tag{17}$$

Nous savons que l'on peut toujours écrire :

$$\Psi_k(x, y, z, t) = \iiint_{-\infty}^{+\infty} \Big[a_k(p_x\,p_y\,p_z)\,\mathrm{e}^{\frac{2\pi i}{h}(Wt - p_x x - p_y y - p_z z)} + b_k(p_x\,p_y\,p_z)\,\mathrm{e}^{\frac{2\pi i}{h}(Wt - p_x x - p_y y - p_z z)} \Big]\,dp_x\,dp_y\,dp_z \tag{18}$$

avec :

$$W = +\sqrt{m_0^2 c^2 + p_x^2 + p_y^2 + p_z^2}. \tag{19}$$

Les huit $a_k(p_x p_y p_z)$ et $b_k(p_x p_y p_z)$ peuvent se calculer à partir de 4 d'entre eux qui sont arbitraires par des formules que nous connaissons.

Appelons maintenant « espace des moments » l'espace constitué en prenant comme coordonnées rectangulaires les grandeurs $p_x\,p_y\,p_z$ et divisons cet espace en cellules σ aussi petites que nous voulons. Les quantités :

$$\Delta(\sigma) = \iiint \mathrm{e}^{-\frac{2\pi i}{h}(p_x x + p_y y + p_z z)}\,dp_x\,dp_y\,dp_z \tag{20}$$

sont (à une constante de normalisation près) les « différen-

(1) *Annales de l'Institut H. Poincaré, loc. cit.*

tielles propres » du spectre continu des ondes monochromatiques planes [1] et nous pouvons écrire :

$$\Psi_k(x, y, z, t) = \sum_\sigma \Big[a_k(p_x\, p_y\, p_z)\, e^{\frac{2\pi i}{h} W t} + b_k(p_x\, p_y\, p_z)\, e^{-\frac{2\pi i}{h} W t} \Big] \Delta(\sigma) \quad (21)$$

$p_x p_y p_z$ étant les coordonnées du centre de l'élément σ dans l'espace des moments et $\sum\limits_\sigma$ désignant la sommation sur toutes les cellules σ de cet espace.

Ceci étant, nous pouvons écrire (17) sous la forme :

$$\frac{d\bar{x}}{dt} = -c \sum_\sigma \sum_{\sigma'} \sum_1^4{}_k \Big[a_k(p'_x\, p'_y\, p'_z)\, e^{\frac{2\pi i}{h} W t} + b_k(p'_x\, p'_y\, p'_z)\, e^{-\frac{2\pi i}{h} W t} \Big]^* \alpha_1 \Big[a_k(p_x\, p_y\, p_z)\, e^{\frac{2\pi i}{h} W t} + b_k(p_x\, p_y\, p_z)\, e^{\frac{-2\pi i}{h} W t} \Big] \,.\, \int_D \Delta^*(\sigma')\, \Delta(\sigma)\, d\tau \quad (22)$$

le domaine étant naturellement ici l'espace tout entier. Les différentielles propres étant orthogonales et pouvant être supposées normées, nous avons [2] :

$$\frac{d\bar{x}}{dt} = -c \sum_\sigma \sigma \sum_1^4{}_k \Big[a_k(p_x\, p_y\, p_z)\, e^{\frac{2\pi i}{h} W t} + b_k(p_x\, p_y\, p_z)\, e^{-\frac{2\pi i}{h} W t} \Big]^* . \alpha_1 . \Big[a_k(p_x\, p_y\, p_z)\, e^{\frac{2\pi i}{h} W t} + b_k(p_x\, p_y\, p_z)\, e^{-\frac{2\pi i}{h} W t} \Big] \quad (23)$$

ou encore :

$$\frac{d\bar{x}}{dt} = -c \sum_\sigma \sigma \sum_1^4{}_k \big[a_k^*\, \alpha_1\, a_k + b_k^*\, \alpha_1\, b_k \big] - c \sum_\sigma \sigma \Big(\sum_1^4{}_k a_k^*\, \alpha_1\, b_k\, e^{-\frac{4\pi i}{h} W t} + \sum_1^4{}_k b_k^*\, \alpha_1\, a_k\, e^{\frac{4\pi i}{h} W t} \Big) \quad (24)$$

---

(1) Voir la définition des différentielles propres au chapitre V, paragr. 4.

(2) Nous désignons ici par σ le volume de la cellule σ.

Il s'agit maintenant de transformer cette expression. Nous avons d'après les formules du chapitre précédent :

$$a_1 = -\frac{p_z A + (p_x - i p_y) B}{\frac{W}{c} + m_0 c} ; \quad a_2 = -\frac{(p_x - i p_y) A - p_z B}{\frac{W}{c} + m_0 c} ;$$

$$a_3 = A ; \quad a_4 = B \tag{25}$$

d'où :

$$-c \sum_{1}^{4}{}_k a_k^* \alpha_1 a_k = -c (a_1^* a_4 + a_2^* a_3 + a_3^* a_2 + a_4^* a_1)$$

$$= 2 p_x c^2 \frac{AA^* + BB^*}{W + m_0 c^2} \tag{26}$$

Mais on a aussi :

$$\sum_{1}^{4}{}_k a_k^* a_k = (AA^* + BB^*)\left(1 + \frac{p_x^2 + p_y^2 + p_z^2}{\left(\frac{W}{c} + m_0 c\right)^2}\right)$$

$$= 2 W \frac{AA^* + BB^*}{W + m_0 c^2} . \tag{27}$$

d'où par comparaison avec (26) :

$$-c \sum_{1}^{4}{}_k a_k^* \alpha_1 a_k = \frac{p_x c^2}{W} \sum_{1}^{4}{}_k a_k^* a_k . \tag{28}$$

On trouverait de même en utilisant l'expression des b_k :

$$-c \sum_{1}^{4}{}_k b_k^* \alpha_1 b_k = -\frac{p_x c^2}{W} \sum_{1}^{4}{}_k b_k^* b_k \tag{29}$$

D'autre part, les deux derniers termes de l'expression (24) sont complexes conjugués (à cause de l'hermiticité de α_1) et peuvent s'écrire :

$$c \sum_{\sigma} A_1 \cos\left(\frac{4\pi}{h} W t + \varphi_1\right) \tag{30}$$

A_1 et φ_1 variant naturellement d'une cellule σ à une autre, c'est-à-dire étant fonctions de $p_x\, p_y\, p_z$.

Finalement par (28), (29) et (30), nous pouvons écrire (24) sous la forme :

$$\frac{d\bar{x}}{dt} = \sum_{\sigma} \sigma \frac{c^2 p_x}{W}\left(\sum_{1}^{4}{}_k a_k^* a_k - \sum_{1}^{4}{}_k b_k^* b_k\right) + \sum_{\sigma} \sigma c A_1 \cos\left(\frac{4\pi W}{h} t + \varphi_1\right) \quad (31)$$

Or, d'après les formules de la Dynamique relativiste, la quantité $\frac{c^2 p_x}{W}$ est la composante x de la vitesse correspondant à la quantité de mouvement p_x et à l'énergie $+W$. De même, on peut considérer $-\frac{c^2 p_x}{W}$ comme la composante x de vitesse correspondant à la quantité de mouvement et à l'énergie négative $-W$. Le premier terme dans l'expression (31) de $\frac{d\bar{x}}{dt}$ est donc une sorte de valeur moyenne de la composante de vitesse v_x correspondant à la composition spectrale (18) de l'onde Ψ. Posons donc :

$$\bar{v}_x = \sum_{\sigma} \sigma \frac{p_x c^2}{W} \sum_{1}^{4}{}_k (a_k^* a_k - b_k^* b_k) \quad (32)$$

$\bar{v}_x$ est naturellement indépendante du temps. Nous avons donc :

$$\frac{d\bar{x}}{dt} = \bar{v}_x + \sum_{\sigma} \sigma c A_1 \cos\left(\frac{4\pi}{h} Wt + \varphi_1\right) \quad (33)$$

d'où par intégration :

$$\bar{x} = C^{te} + \bar{v}_x t + \sum_{\sigma} \sigma \frac{hc A_1}{4\pi W} \sin\left(\frac{4\pi}{h} Wt + \varphi_1\right) \quad (34)$$

Posons :

$$\xi = C^{te} + \bar{v}_x t \quad (35)$$

Il vient :

$$\bar{x} = \xi + \sum_{\sigma} \sigma \frac{hc A_1}{4\pi W} \sin\left[2\pi \cdot \frac{2W}{h} \cdot t + \varphi_1\right] \quad (36)$$

On trouverait de même :

$$\bar{y} = \eta + \sum_{\sigma} \sigma \frac{hc A_2}{4\pi W} \sin\left[2\pi \cdot \frac{2W}{h} t + \varphi_2\right];$$

$$\bar{z} = \zeta + \sum_{\sigma} \sigma \frac{hc A_3}{4\pi W} \sin\left[2\pi \cdot \frac{2W}{h} t + \varphi_3\right]. \quad (37)$$

avec les définitions :

$$\eta = \overline{v_y}\, t + C^{te} \quad ; \quad \overline{v_y} = \sum_\sigma \sigma \frac{p_y c^2}{W} \cdot \sum_1^4 {}_k (a_k^* a_k - b_k^* b_k)$$

$$\zeta = \overline{v_z}\, t + C^{te} \quad ; \quad \overline{v_z} = \sum_\sigma \sigma \frac{p_z c^2}{W} \cdot \sum_1^4 {}_k (a_k^* a_k - b_k^* b_k) \tag{38}$$

Le point de coordonnées ξ, η, ζ se déplace d'un mouvement rectiligne et uniforme, mais le centre de gravité de la probabilité exécute autour de ce point une série d'oscillations de fréquences $\frac{2\,W}{h}$. C'est le « tremblement de Schrödinger ». Les amplitudes de ces oscillations sont d'ailleurs en général faibles car elles sont proportionnelles au facteur $\frac{hc}{4\pi W}$ qui est toujours plus petit que $\frac{hc}{4\pi m_0 c^2} = \frac{1}{4\pi} \cdot \frac{h}{m_0 c}$; or la quantité $\frac{h}{m_0 c}$ souvent appelée « longueur d'onde de Compton » est très petite ($2{,}4 \,.\, 10^{-10}$ cm.).

L'analyse précédente montre très nettement l'origine du tremblement de Schrödinger. Il est dû au *battement* des ondes à énergie positive W avec les ondes à énergies négatives $-$ W correspondantes. La fréquence du battement est comme à l'ordinaire la différence de fréquence des ondes battantes, soit ici $\frac{2\,W}{h}$.

Pour un train d'ondes dans la décomposition spectrale duquel ne figure aucune onde à énergie négative, il n'y a pas de tremblement de Schrödinger et le théorème d'Ehrenfest est valable. Mais nous savons que dans le cas général, pour représenter un train d'ondes, il faut faire appel aux ondes d'énergie négative. Voilà pourquoi le théorème n'est pas en général valable dans la théorie de Dirac.

Le tremblement de Schrödinger et la non-validité du théorème d'Ehrenfest sont donc liés à l'existence des états d'énergie négative et disparaîtraient avec eux si on pouvait les éliminer.

CHAPITRE XXII

Quelques remarques sur la Relativité et la nouvelle Mécanique

1. *L'absence de symétrie entre espace et temps en nouvelle Mécanique*

NOUS avons vu que la théorie de Dirac est sous certains aspects conforme au principe de Relativité. On peut, en effet, conserver à ses équations fondamentales une forme qui est invariante pour une transformation de Lorentz : on peut former à l'aide de ses 4 Ψ des grandeurs qui ont un caractère tensoriel dans l'espace-temps. Néanmoins on ne peut guère prétendre que la théorie de Dirac, dans son état actuel, soit complètement en accord avec les conceptions de la Relativité, même sous sa forme restreinte. Une des idées directrices de la théorie de Relativité paraît être, en effet, de toujours faire intervenir d'une façon symétrique les coordonnées d'espace et de temps. Or, en théorie de Dirac, cette intervention symétrique des variables $xyzt$ n'est pas réalisée car on y admet les principes généraux de la nouvelle Mécanique, principes généraux qui, du moins sous leur forme présente, donne à la variable « temps » un rôle tout à fait particulier. Nous devons insister sur ce point.

D'abord la nouvelle Mécanique fait correspondre à toute grandeur physique observable un opérateur hermitique. Or

l'hermiticité d'un opérateur est définie dans un domaine *d'espace* et cela suffit déjà pour que la définition même des opérateurs employés par la nouvelle Mécanique ne soit pas relativiste. Le temps ne peut figurer dans ces opérateurs qu'à titre de paramètre et les dérivées $\frac{\partial^n}{\partial t^n}$ ne peuvent jamais y figurer.

Ayant défini les opérateurs qui correspondent aux grandeurs observables, la nouvelle Mécanique admet que les valeurs possibles d'une de ces grandeurs sont données par les valeurs propres de son opérateur. Mais les valeurs propres et les fonctions propres d'un opérateur sont à leur tour définies dans un domaine D de l'espace. La variable « temps » ne joue aucun rôle dans le calcul des valeurs et fonctions propres d'un opérateur hermitique ; quand elle y figure, c'est simplement à titre de paramètre. Les valeurs possibles d'une grandeur observable A étant ainsi calculées, la nouvelle Mécanique admet que les probabilités respectives des diverses valeurs possibles de cette grandeur sont données par le carré des modules du coefficient de chaque fonction propre dans le développement de la fonction d'ondes Ψ suivant ces fonctions propres.

Dans le cas général (1), ces probabilités dépendent du paramètre t et c'est précisément pour cela qu'en général l'état du système évolue.

Bref, la nouvelle Mécanique considère jusqu'à présent le temps comme jouant un rôle tout différent de celui des coordonnées d'espace. La théorie de Dirac qui admet les principes généraux de la nouvelle Mécanique, ne peut donc pas être réellement « relativiste ». Ceci se voit aisément par exemple quand on étudie les valeurs moyennes des grandeurs en Mécanique de Dirac. Les densités de valeur moyenne sont définies dans l'espace et pour passer de la densité à la valeur moyenne elle-même, il faut faire une intégration dans l'es-

(1) C'est-à-dire quand A n'est pas intégrale première

pace, opération qui n'est pas relativistiquement invariante. On peut ainsi expliquer certaines particularités très visibles sur le tableau des valeurs moyennes qui se trouve à la fin de la deuxième partie de ce livre, comme, par exemple, l'intervention très dissymétrique de l'opérateur α_1.

2. ***La quatrième relation d'incertitude.*** — Nous allons signaler encore une question qui est en relation avec le rôle particulier joué par le temps en Mécanique ondulatoire.

Les relations d'incertitude d'Heisenberg sont les suivantes

$$\Delta p_x \cdot \Delta x \geqslant h \; ; \quad \Delta p_y \cdot \Delta y \geqslant h \; ; \quad \Delta p_z \cdot \Delta z \geqslant h \, . \qquad (1)$$

Au point de vue mathématique, elles dérivent du fait que la Mécanique ondulatoire fait correspondre (à une constante près) aux grandeurs classiques $p_x \, p_y \, p_z$ les dérivées par rapport aux variables conjuguées x, y, z respectivement.

Au point de vue physique, les relations (1) doivent s'interpréter en disant : à un instant donné, le produit de l'incertitude sur une des coordonnées d'un corpuscule par l'incertitude sur la composante conjuguée de quantité de mouvement est toujours au moins de l'ordre de h.

La symétrie relativiste entre espace et temps exigerait que les trois relations (1) soient complétées par la quatrième relation :

$$\Delta W \, . \, \Delta t \geqslant h \qquad (2)$$

puisque l'énergie W est la composante de temps d'un quadrivecteur « Impulsion d'Univers » dont p_x, p_y, p_z sont les composantes d'espace.

Mais, dans l'état actuel de la nouvelle Mécanique, cette quatrième relation d'incertitude ne peut pas du tout être interprétée de même que les trois premières car, d'une part, le temps t doit être considéré comme un paramètre ayant une valeur bien déterminée sans aucune incertitude et, d'autre part, la grandeur « énergie W » correspond à l'opérateur Hamiltonien et non pas à l'opérateur $\dfrac{h}{2\pi i} \dfrac{\partial}{\partial t}$ qui ne peut pas être considéré comme hermitique au sens propre du mot.

Il est néanmoins possible de donner un sens à l'égalité (2). En effet, il est bien connu que si l'on observe le passage d'une onde en un point fixe de l'espace pendant un temps fini quelconque Δt, on ne peut affirmer que cette onde possède la fréquence ν qu'avec une incertitude :

$$\Delta\nu \geqslant \frac{1}{\Delta t} \tag{3}$$

Pour une onde Ψ en tenant compte de la relation $W = h\nu$, on peut écrire l'inégalité (3) sous la forme :

$$\Delta W \geqslant \frac{h}{\Delta t}. \tag{4}$$

On aperçoit alors le sens de la relation (2). Elle exprime qu'une expérience ou observation faite en un point fixe pendant une durée Δt, ne peut pas faire connaître l'énergie d'un corpuscule avec une incertitude moindre que $\frac{h}{\Delta t}$.

La quatrième relation d'incertitude a donc un sens, mais un sens très différent de celui des trois premières. C'est là un nouvel aspect de la dissymétrie entre espace et temps en Mécanique ondulatoire.

3. ***Pourrait-on dans la nouvelle Mécanique rétablir la symétrie entre espace et temps?*** — Pour rétablir la symétrie entre l'espace et le temps dans la nouvelle Mécanique, il faudrait essayer de modifier dans ce sens l'énoncé des principes généraux. Sans affirmer que cela soit impossible, nous allons montrer qu'il y a à cela de grandes difficultés.

La première chose à faire serait de définir l'hermiticité des opérateurs dans l'espace-temps au lieu de la définir dans l'espace. Soit par exemple un opérateur A pouvant opérer à la fois sur les variables de temps et d'espace (et aussi éventuellement sur la variable de spin ζ). On aurait alors à définir l'hermiticité par la condition :

$$\iiiint_D f^*(x, y, z, t)\, A\, g(x, y, z, t)\, dx\, dy\, dz\, dt$$
$$= \iiiint_D g(x, y, z, t)\, A^* f^*(x, y, z, t)\, dx\, dy\, dz\, dt. \tag{5}$$

D étant un domaine d'espace-temps, pour l'électron libre l'espace-temps tout entier.

Eventuellement, il y aurait lieu d'ajouter dans la condition (5) une sommation sur la variable ζ.

Les valeurs propres et les fonctions propres de l'opérateur A seraient alors définies par l'équation :

$$A\left[\varphi_i(x, y, z, t)\right] = \alpha_i \varphi_i(x, y, z, t) \tag{6}$$

φ_i étant finie, uniforme, continue et nulle aux limites du domaine D d'espace-temps. La fonction d'onde $\Psi(x, y, z, t)$ se développerait sous la forme :

$$\Psi(x, y, z, t) = \sum_i c_i \varphi_i(x, y, z, t) \tag{7}$$

et l'on pourrait transposer l'énoncé habituel des principes généraux en disant : « une mesure de la grandeur physique correspondant à l'opérateur A ne peut donner qu'une des valeurs α_i et la probabilité de la valeur α_k est égale à $|c_k|^2$ ». De là résulterait que la valeur moyenne de la grandeur A serait :

$$\overline{A} = \int\int\int\int_D \Psi^* A(\Psi)\, dx\, dy\, dz\, dt. \tag{8}$$

Malheureusement, il y a à tout cela une objection grave. Les quantités α_i et c_i seraient, d'après leurs nouvelles définitions, indépendantes du temps et naturellement il en serait de même pour $\overline{A}$. Nous obtiendrions ainsi une Physique statique d'où toute évolution dans le temps serait bannie.

Un aspect différent de la difficulté est le suivant. Quand, avec les principes actuels de la nouvelle Mécanique, nous « quantifions » un système, un atome d'hydrogène par exemple, nous isolons par la pensée ce système du reste de l'univers. Rigoureusement parlant, ceci n'est pas permis : pour déterminer le Ψ, il faudrait en principe tenir compte non seulement du champ de force créé par le noyau, mais de tous les champs de force existant dans l'univers entier. Fort

heureusement, l'influence des champs de force extérieurs à l'atome sur la forme des ondes Ψ stationnaires dans l'atome est tout à fait négligeable parce que ces ondes Ψ tendent très rapidement vers zéro quand on s'éloigne du domaine atomique. En principe, la détermination des ondes stationnaires, des fonctions propres, exige la considération de l'espace entier et de tout ce qui y est contenu, mais en pratique la structure du monde matériel se prête à ce que nous y découpions des systèmes suffisamment indépendants du reste du monde matériel pour que nous puissions les considérer isolément. Mais si nous voulons définir les fonctions propres dans l'espace-temps, il n'en sera plus de même car il ne paraît pas possible de découper l'existence d'un individu physique tel qu'un atome en sections indépendantes les unes des autres. Considérons un atome d'hydrogène. Au cours de son histoire, il subira des actions diverses, sera par exemple le siège d'effets Stark ou d'effets Zeeman. Si nous voulons définir les fonctions et les valeurs propres dans l'espace-temps, nous trouverons que les états stationnaires de cet atome seront invariables et seront déterminés par l'ensemble de toutes les actions qu'il subit pendant tout le cours du temps. Ceci ne paraît guère acceptable.

En réalité, même dans la théorie de Relativité sous sa forme aujourd'hui classique, les variables de temps et d'espace sont loin d'être équivalentes. La variable « temps » y varie toujours dans le même sens et les lignes d'Univers de tous les objets matériels sont des lignes ayant un sens positif, sens positif qui fait toujours un angle de moins de 45° avec l'axe *ct*. Autrement dit, l'espace-temps possède une « polarité » essentielle.

Dans la conception relativiste, un observateur A considère comme simultanés et correspondant à une même valeur de son temps propre les points d'Univers qui sont contenus dans une certaine section à trois dimensions de l'espace-temps ; c'est parce que cette section coupe toutes les lignes d'Univers que l'observateur A peut découper dans son espace

des unités presque indépendantes. Mais dans le sens des lignes d'Univers, un tel découpement serait impossible. Il y a une sorte de structure fibreuse de l'espace-temps dans le sens du temps ; c'est cette structure fibreuse qui nous gène ici et nous voyons que la difficulté a sa racine dans la Relativité classique elle-même.

4. ***Forme plus restrictive des relations d'incertitude*** (***Bohr, Landau et Peïerls***) — La relation $\Delta W . \Delta t \geqslant h$ conduit, si l'on introduit l'idée relativiste qu'aucune action ne peut se propager avec une vitesse supérieure à c, à énoncer de nouvelles relations d'incertitude ; ces nouvelles relations qui ne sont pas contenues dans les principes généraux de la nouvelle Mécanique sous sa forme non-relativiste, viennent s'ajouter à celles d'Heisenberg et augmenter les incertitudes qui en résultent.

Sans entrer dans le détail des raisonnements qui conduisent à ces nouvelles incertitudes, nous allons cependant chercher à en montrer l'origine. Plaçons-nous dans le cas de l'absence de champ. En Mécanique ondulatoire relativiste, les 4 paramètres W, p_x, p_y, p_z qui figurent dans la phase d'une onde plane monochromatique sont reliés par la relation de Dynamique Einsteinienne :

$$|p|^2 = \frac{W^2}{c^2} - m_0^2 c^2 \tag{9}$$

d'où résulte aisément $\Delta |p| \geqslant \frac{\Delta W}{c}$.

Mais nous savons que si l'on cherche à déterminer l'état d'un électron par une expérience de durée Δt, on ne peut connaître qu'avec une incertitude $\frac{h}{\Delta t}$ c'est-à-dire que la décomposition de l'onde Ψ en ondes monochromatiques intéresse un intervalle spectral :

$$\Delta \nu = \frac{\Delta W}{h} \geqslant \frac{1}{\Delta t} . \tag{10}$$

Les valeurs de $|p|$ qui interviennent dans la décomposition spectrale de l'onde Ψ occupent donc un domaine :

$$\Delta\,|p\,| \geqslant \frac{\Delta\,\mathrm{W}}{c} \geqslant \frac{h}{c\,\Delta\,t}\,. \tag{11}$$

On en conclut qu'une observation ou expérience de durée Δt ne peut pas conduire à connaître la quantité de mouvement avec une incertitude moindre que $\frac{h}{c\Delta t}$. Ceci fournit une relation d'incertitude d'un caractère nouveau.

On peut encore retrouver l'inégalité (11) d'une autre manière. Si un observateur placé en un point de l'espace veut délimiter la longueur d'un train d'ondes Ψ, il doit le faire passer au travers d'une ouverture percée dans un écran et pouvant s'obturer au moyen d'un volet. Il enlèvera le volet pendant un intervalle de temps Δt, puis le remettra en place de façon à délimiter le train d'ondes ayant traversé l'écran pendant ce temps Δt. Mais le front de ce train d'ondes ne pouvant pas progresser avec une vitesse supérieure à c, la longueur du train d'ondes transmis à travers l'écran est au plus égale à $c\Delta t$. En prenant l'axe des z dans la direction de propagation, l'incertitude sur la coordonnée z du corpuscule dans l'onde transmise sera donc :

$$\Delta z \leqslant c\Delta t \tag{12}$$

et, d'après la relation d'Heisenberg :

$$\Delta\,p_z \geqslant \frac{h}{\Delta z} \geqslant \frac{h}{c\,\Delta\,t} \tag{13}$$

et comme ici on peut confondre p et p_z, on retrouve l'inégalité (11).

Au premier abord, il pourrait sembler que la relation (11) est en opposition avec la notion d'onde plane monochromatique car, pour celle-ci, on a $\Delta p = 0$. Mais on peut écarter cette objection en remarquant que, pour être rigoureusement en droit d'associer à un corpuscule une onde Ψ plane et monochromatique, il faudrait pouvoir affirmer que le

corpuscule peut se trouver en n'importe quel point de l'espace et cette affirmation ne pourrait être justifiée que pour une expérience d'une durée infinie puisqu'aucun moyen d'investigation ne peut explorer l'espace avec une vitesse supérieure à c. Il reste cependant de cette objection qu'en toute rigueur, un observateur ne peut jamais représenter ses connaissances sur l'état d'un électron par une onde plane monochromatique illimitée.

Certains raisonnements portent à croire que pour un corps de masse propre m_0, la mesure d'une longueur inférieure à $\frac{h}{m_0c}$ ou d'une durée inférieure à $\frac{h}{m_0 c^2}$ est illusoire (1). Ceci permet d'espérer que si l'on pouvait exclure de la théorie de l'électron la considération des distances inférieures à $\frac{h}{m_0c}$ et des intervalles de temps inférieurs à $\frac{h}{m_0c^2}$ comme étant dépourvue de sens, on parviendrait peut-être par là-même à faire disparaître la difficulté des énergies négatives. Mais ce n'est là encore qu'un espoir.

(1) Cf. Schrödinger, *Annales de l'Institut H. Poincaré, loc. cit.*

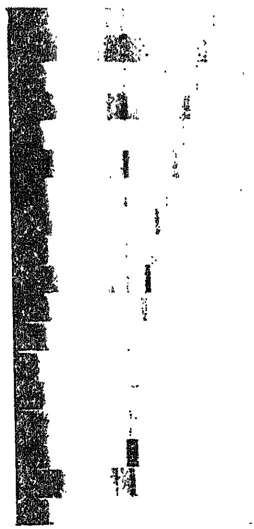


FIG. 1

Bandes d'absorption K des éléments
du numéro 55 au numéro 73

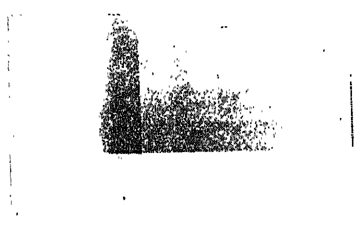


FIG. 2

Bandes d'absorption L de l'Uranium
et du Thorium

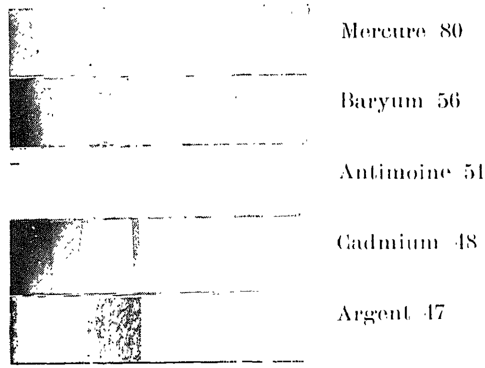


Fig. 3

Discontinuités K d'absorption au voisinage de la bande de l'argent

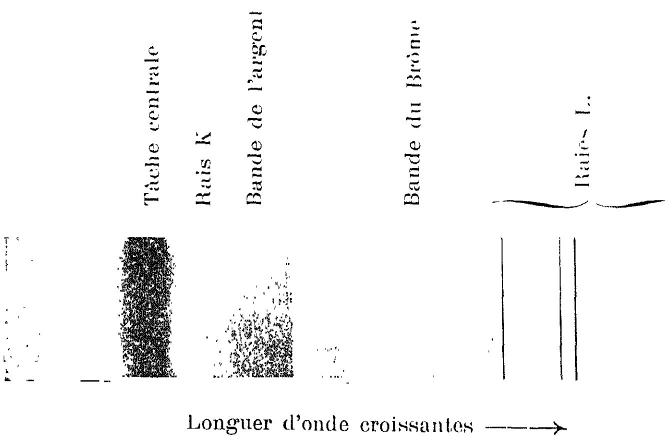


Fig. 4

Spectre d'émission K et L du Tungstene et bandes d'absorption renversées du brôme et de l'argent de l'émulsion photographique

TABLE DES MATIÈRES

PREMIÈRE PARTIE

Succès et échecs des théories quantiques et de la mécanique ondulatoire primitive

CHAPITRE PREMIER

Le spectre atomique de l'hydrogène Théories de Bohr et Sommerfeld

CHAPITRE II

Notions sommaires sur les spectres optiques de doublets et leur interprétation

CHAPITRE III

Les spectres de rayons X et les théories de Bohr et Sommerfeld

Pages

CHAPITRE IV

Les anomalies magnétiques et l'hypothèse de l'électron tournant

CHAPITRE V

Résumé des principes de la mécanique ondulatoire

CHAPITRE VI

Résumé des principes de la mécanique ondulatoire *(suite)*

CHAPITRE VII

Forme relativiste de la mécanique ondulatoire à une fonction d'onde

CHAPITRE VIII

Succès et échecs de la mécanique ondulatoire à une fonction d'onde

Pages

DEUXIEME PARTIE

La théorie de l'électron magnétique et tournant de Dirac
Principes généraux

CHAPITRE IX

La tentative de Pauli

CHAPITRE X

La théorie de Dirac

CHAPITRE XI

Invariance relativiste des équations de Dirac

CHAPITRE XII

Densité et courant dans la théorie de Dirac
Ondes planes

CHAPITRE XIII

Le magnétisme propre de l'électron

Pages

CHAPITRE XIV

Le tenseur « densité de moment électrique et de moment magnétique »

CHAPITRE XV

Matrices et intégrales premières dans la théorie de Dirac Moment de rotation propre de l'électron

CHAPITRE XVI

Résumé systématique des résultats obtenus

TROISIÈME PARTIE

Applications de la théorie de Dirac Critiques et compléments divers

CHAPITRE XVII

Explication de la structure fine par la théorie de Dirac

Imp. G. Thone, Liège (Belgique) 10-33